PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/47, C12N 15/62, C07K 16/18, A61K 38/16, G01N 33/68, A01K 67/027, C12N 15/11, A61K 48/00

A2

(11) International Publication Number: **WO 99/41376**

(43) International Publication Date: 19 August 1999 (19.08.99)

(21) International Application Number: PCT/US99/03072

(22) International Filing Date: 12 February 1999 (12.02.99)

(30) Priority Data:

60/074,559	12 February 1998 (12.02.98)	US
Not furnished	11 February 1999 (11.02.99)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application

US	60/074,559 (CIP)
Filed on	12 February 1998 (12.02.98)

(71) Applicant *(for all designated States except US)*: CURAGEN CORPORATION [US/US]; 11th floor, 555 Long Wharf Drive, New Haven, CT 06511 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: YANG, Meijia [US/US]; 6 Catbird Lane, East Lyme, CT 06333 (US). NANDABALAN, Krishnan [US/US]; 228 Village Pond Road, Guilford, CT 06437 (US). SCHULZ, Vincent, Peter [US/US]; 21 Old Farms Road, Madison, CT 06443 (US).

(74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Cohn, Ferris, Glovsky, and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: RETINOBLASTOMA PROTEIN COMPLEXES AND RETINOBLASTOMA INTERACTING PROTEINS

(57) Abstract

The present invention relates to complexes of the retinoblastoma (RB) protein with proteins identified as interacting with RB by an improved, modified yeast two hybrid assay system. The proteins identified to interact with RB are RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392*. Thus, the present invention provides complexes of RB and RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392* and derivatives, fragments and analogs thereof. The invention also provides the *CD24**, *Set**, *GluT1**, and *115392** genes and proteins and derivatives, fragments and analogs thereof. Methods of screening the complexes for efficacy in treating and/of preventing certain diseases and disorders, particularly hyperproliferative disorders, including cancer, neurodegenerative disease, and viral disease, are also disclosed herein.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova.	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

RETINOBLASTOMA PROTEIN COMPLEXES AND RETINOBLASTOMA INTERACTING PROTEINS

RELATED PATENT APPLICATIONS AND GRANT SUPPORT

The instant application claims priority to Provisional Application Serial No. 60/074559 entitled, "RETINOBLASTOMA PROTEIN COMPLEXES", filed February 12, 1998.

This invention was made with United States Government support under award number 70NANB5H1066 awarded by the National Institute of Standards and Technology. The United States Government has certain rights in the invention.

FIELD OF THE INVENTION

The present invention discloses complexes of the retinoblastoma protein (RB) with other proteins, in particular, complexes of RB with RN-tre, RB with the cellular apoptosis susceptibility protein (CAS), RB with the gamma-interferon inducible protein (IP-30), RB with the ribosomal protein L6, RB with the cytokine IK factor, RB with lactate dehydrogenase B (LDH-B), RB with Nlk1(Nek 2), RB with cyclophilin A, RB with Zap 3, RB with the oxidoreductase 17-beta-hydroxysteroid (17 beta-HSD6), RB with CD24*, RB with Set*, RB with the glutamate transporter GluT1*, and RB with clone 115392*. The present invention also discloses antibodies to RB complexes, and their use in, *inter alia,* screening, diagnosis, prognosis and therapy. The present invention further discloses CD24*, Set*, GluT1*, and 115392* genes and proteins, as well as derivatives, fragments and analogs, thereof.

BACKGROUND OF THE INVENTION

Human Retinoblastoma Protein (RB) complexes with one or more of the following proteins: RN-tre, cellular apoptosis susceptibility protein CAS, gamma-interferon inducible protein IP-30, ribosomal protein L6, cytokine IK factor, lactate dehydrogenase B (LDH-B), Nlk1(Nek 2), cyclophilin A, Zap 3, oxidoreductase 17-beta-hydroxysteroid (17 beta-HSD6), CD24*, Set*, glutamate transporter GluT1* or clone 115392*, have not been previously described prior to the disclosure of the present invention.

It should be noted that the citation of a reference herein shall not be construed as an admission that such is prior art to the present invention.

(I) Human Retinoblastoma Protein (RB)

The human retinoblastoma tumor suppressor protein (hereinafter designated "RB"; GenBank Acc. No. M28419, U.S. Patent 4,942123; Lee, *et al*, 1987. *Nature* 329: 642-645) is considered a classical tumor suppressor protein. The RB protein belongs to the retinoblastoma "tumor suppressor family" that includes RB, pRb2/p130, and p107. Members of the retinoblastoma family are involved in implementing and controlling three major aspects of cellular life: (*i*) proliferative growth, (*ii*) differentiation, and (*iii*) apoptosis (see *e.g.*, Stiegler, 1998. *J. Cell. Biochem.* Suppl. 30-31:30-36). The activities of these proteins are highly regulated, enabling them to precisely establish control. Proteins that interact with RB have potentially wide ranging effects on both normal physiological processes and important pathogenic processes.

The phosphoprotein is the critical control protein for cellular passage through the G1 restriction point of the cell cycle (see *e.g.*, Weinberg, 1995. *Cell* 81:323-330). The overexpression of RB, and maintenance of RB in its hypo-phosphorylated form, results in inhibition of the cell cycle. In contrast, RB needs to be phosphorylated by cyclin-dependent kinases before mammalian cells can continue the cell cycle.

Growth suppression by RB is dependent on its ability to form complexes with transcription regulators. At least three distinct protein-binding activities have been identified in RB: the large A/B pocket binds E2F, the A/B pocket binds the LXCXE peptide motif (*e.g.*, insulin, papilloma virus E6 and E7 proteins), and the C pocket binds the nuclear c-Abl tyrosine kinase (see *e.g.*, Whitaker, *et al.*, 1998. *Mol. Cell. Biol.* 18:4032-4042). Retinoblastoma protein binds to several proteins involved in transcription, including but not limited to the (i) E2F, (*ii*) the transcription factor UBF (see *e.g.*, Cavanaugh, *et al.*, 1995. *Nature* 374:177-180); (iii) the transcription regulator Brm protein (see *e.g.*, Singh, *et al.*, 1995. *Nature* 374:562-565); (*iv*) the transcriptional regulator RIZ (see *e.g.*, Buyse, *et al.*, 1995. *Proc. Natl. Acad. Sci. USA* 92:4467-4471); (*v*) the transcriptional activation protein IE2 (see *e.g.*, Choi, *et al.*, 1995. *Virology* 208:450-456); (v) the transcriptional effectors Elf-1, MyoD, PU.1, ATF-2 (see *e.g.*, Wang, *et al.*, 1994. *Adv. Cancer Res.* 64:25-85). The signaling molecule Raf-1 can physically interact with RB and thus, this complex may play a role in the growth factor-mediated induction of cell proliferation and provide a link between cell surface signaling cascades and the cell cycle machinery. See *e.g.*, Wang, *et al.*, 1998. *Mol. Cell. Biol.* 18:7487-98.

RB is implicated in the pathogenesis of a number of cancers, including, but not limited to cervical cancer, breast cancer, retinoblastomas, pituitary tumors, lymphoma, small cell carcinoma of the lung, esophageal cancer, glioblastoma, familial melanoma, pancreatic cancer, sarcoma and bladder cancer. The mechanism of RB-mediated carcinogenesis has been elucidated for a number

of cancers due to its central role in the control of cell cycle progression. For example, in the case of most cervical cancers, the human papilloma virus E7 oncoprotein has been demonstrated to bind to and inactivate RB. See *e.g.*, Dyson, *et al.*, 1989. *Science* 243:934-937. Similarly, in many squamous cell esophageal, breast, and B cell lymphomas, the over-expression of cyclin D causes the inactivation of RB. See *e.g.*, Jiang, *et al.*, 1992. *Cancer Res.* 52:2980-2983. A mechanism which has been implicated in esophageal, lung, bladder, and pancreatic cancers, familial melanomas and glioblastomas appears to involve the RB inactivation by cyclin dependend kinases. See *e.g.*, Weinberg, 1995. *Cell* 81:323-330. In addition, RB has been demonstrated to bind to, and is concomitantly down-regulated by, the cellular oncoprotein MDM2 in human sarcomas, gliomas and squamous cell carcinomas. See *e.g.*, Xiao, *et al.*, 1995. *Nature* 375:694-698. Finally, RB has been demonstrated to prevent the interaction of cellular and viral oncoproteins (such as E1A and c-Myc) with a transcription factor. See *e.g.*, Hateboer, *et al.*, 1993. *Proc. Natl. Acad. Sci. USA* 90:8489-8493. Disruption of normal cell-to-cell contact inhibition mediated by RB also appears to play a role in the metastatic behavior of tumor cells. See *e.g.*, Zhu, *et al.*, 1996. *J. Cell Biol.* 133:391-403.

In addition to the function as tumor suppressor, RB displays an independent role in the differentiation of some cell types. For example, the nerve growth factor-mediated accumulation of hypophosphorylated RB is an important signal for differentiation of cells. See *e.g.*, Li, *et al.*, 1996. *J. Neurochem.* 66:2287-2294.

Thus, in summary, RB has been centrally implicated in numerous physiological processes, including, but not limited to: (*i*) proliferative growth, (*ii*) differentiation, and (*iii*) apoptosis. In addition, RB has been strongly implicated in play a role in numerous pathophysiological conditions, including, but not limited to: (*i*) hyperproliferative states (*e.g.*, tumorigenesis and metastatic tumor dissemination); (*ii*) degenerative states (*e.g.*, neurodegeneration) and (*iii*) response to viral infection.

(II) RN-tre

The protein designated "Related to the -N-terminus of tre" (hereinafter designated RN-tre; submitted as GenBank Acc. Nos. D13644; Nomura, *et al.*, 1994. *DNA Res.* 1:27-35) is a 97-100 kD protein which appears to have been derived from a fusion of an amino-terminal tre-like region to a carboxy-terminal region encoding a deubiquinating enzyme. See *e.g.*, Matoskova, *et al.*, 1996. *Oncogene* 12:2653-2671. The *RN-tre* gene maps to chromosomal location 10p13, a region known to be a site of translocations in various leukemias and a monosomy syndrome characterized by developmental anomalies. The tre domain within the N-terminus is conserved among several proteins in species from yeast to humans and has demonstrable protein-binding properties. The best known substrate for RN-tre is eps8, a protein involved in mitogenic signaling. See *e.g.*, Biesova, *et al.*, 1997. *Oncogene* 14:233-241

(III) Cellular Apoptosis Susceptibility Protein (CAS)

The human chromosome segregation gene homolog (hereinafter designated *CAS*; GenBank Acc. No. U33286; Brinkmann, *et al.*, 1995. *Proc. Natl. Acad. Sci. USA* 92:10427-10431) is a cellular apoptosis susceptibility gene. CAS plays a critical role in toxin- and tumor necrosis factor alpha 1-mediated cell death. See *e.g.*, Brinkmann, *et al.*, 1996. *Biochemistry* 35:6891-6899. The protein is found to be highly expressed in human leukemia, colon cancer, and breast cancer cell lines, as well as human testis and fetal liver, all of which contain actively dividing cells, supporting a role for the protein in either cell proliferation, or in a homeostatic response to cell proliferation. The *CAS* gene has been mapped to chromosome 20q13.2, a region that harbors mutations associated with aggressive breast tumors. See *e.g.*, Brinkmann, *et al.*, 1996. *Genome Res.* 6:187-194. CAS associates with microtubules and the mitotic spindle (see *e.g.*, Scherf, *et al.*, 1996. *Proc. Natl. Acad. Sci. USA* 93:2670-2674), thus it may impact normal cellular mitosis, cellular apoptosis, and carcinogenesis.

(IV) Human Gamma-Interferon-Inducible Protein (IP-30)

The human gamma-interferon-inducible protein IP-30 (GenBank Acc. No. J03909; Luster, *et al.*, 1988. *J. Biol. Chem.* 263:12036-12043) is induced by interferon gamma which is believed to orchestrate cellular response to various pathogens. See *e.g.*, Boehm, *et al.*, 1997. *Ann. Rev. Immunol.* 15:749-795. Immunohistochemical studies indicate a possible lysosomal localization for IP-30, supporting a role in protein breakdown. The exact function of IP-30 is not yet known, but it may play a role in gamma-interferon mediated immune reactions.

(V) Human Ribosomal Protein (L6)

The human ribosomal protein L6 (GenBank Acc. No. X69391; Zaman, 1993. *Nucl. Acids Res.* 21:1673) is the human homolog of the rat L6 ribosomal protein, which is a structural component of the large ribosomal subunit, and may serve to regulate protein translation (see *e.g.*, Chan, *et al.*, 1996. *Biochem. Mol. Biol. Int.* 39:431-438).

(VI) Human IK Factor

The human IK factor (GenBank Acc. No. S74221; Krief, *et al.*, 1994. *Oncogene* 9:3449-3456; U.S. Patent No. 5,612,195; PCT Publication WO 93/11232; GenBank Acc. No. A25270) is a cytokine which has been shown to be secreted by a number of cancer cell lines. IK inhibits gamma-interferon induction in normal as well as leukemic and other cancer cell lines. The factor is also believed to play a role in the escape of cancer cells from immunorecognition.

(VII) Human Lactate Dehydrogenase B (LDH-B)

Human lactate dehydrogenase B (hereinafter designated LDH-B; GenBank Acc. No. Y00711; Sakai, *et al.*, 1987. *Biochem. J.* 248:933-936) is the B subunit of lactate dehydrogenase, which is widely used as a clinical marker of liver, pancreas, and gall bladder function. It is associated with germ cell tumors and is a valuable prognostic marker in lymphoma, leukemia, neuroblastoma and colon cancer. See *e.g.*, Brodeur, *et al.*, 1993. *J. Clin. Oncol.* 11:1466-1477.

(VIII) Nlk-1(Nek2)

Nlk1 (also known as Nek 2; GenBank Acc. No. U11050) is the human homolog of the mitotic regulator NIMA of *Aspergillus nidulans*. The serine/threonine specific kinase Nlk1(hereinafter designated Nlk1(Nek 2)) has been shown to exhibit a 47% identity to NIMA within its catalytic domain. See *e.g.*, Fry, *et al.*, 1995. *J. Biol. Chem.* 270:12899-12905. Nlk1(Nek 2) appears to function similarly to NIMA. See *e.g.*, Fry and Nigg, 1995. *Curr. Biol.* 5:1122-1125. However, Nlk1(Nek 2) differs from NIMA in that Nlk1(Nek 2) activity peaks in the S/G2 phase of the cell cycle and not in mitosis, and Nlk1(Nek 2) lacks CDC2 phosphorylation sites and C-terminal PEST sequences found in NIMA. Accordingly, the exact function of Nlk1(Nek 2) remains unknown, although, as with other members of the Nlk1(Nek) family, there is strong evidence that it functions in cell cycle control.

(IX) Cyclophilin A

Cyclophilin A (GenBank Acc. No. Y00052; Haendler, *et al.*, 1987. *EMBO J.* 6:947-950), is a widely expressed, small, soluble protein that is the target of cyclosporin A binding. Cyclophilin A functions as a chemo-attractant for macrophages and eosinophils which promote the inflammatory response. Cyclophilin A has also been associated with systemic lupus erythematosus, Lyme disease and rheumatoid arthritis.

Another major function of cyclophilin A, when bound by cyclosporin, is inhibition of calcineurin phosphatase activity, thus affecting a number of cell activities, such as dephosphorylation of transcription factors to prevent their import into the nucleus. See *e.g.*, LeHir, *et al.*, 1995. *Lab Invest.* 73:727-733. Similarly, cyclophilin A itself is targeted to the nucleus, where it can act as a DNase involved in cell apoptosis. See *e.g.*, Montague, *et al.*, 1997. *J. Biol. Chem.* 272:6677-6684. Finally, cyclophilin A has also been demonstrated to specifically associate with the HIV-1 gag polyprotein p55gag, but not other primate immunodeficiency virus coat proteins. This association appears to be necessary for replication of the virus. See *e.g.*, Franke, *et al.*, 1994. *Nature* 372:359-362; Thali, *et al.*, 1994. *Nature* 372:363-365.

(X) Human Zap 3 Protein

The human *Zap 3* gene (GenBank Acc. No. L40403; Sherrington, *et al.*, 1995. *Nature* 375:754-760) is located in the early-onset Alzheimer's disease-associated AD3 locus at 14q24.3. While it is proline-rich, similar to many secreted proteins, the function of the protein sequence has yet to be quantitatively elucidated.

(XI) Human Oxidoreducatse

Human oxidoreductase (GenBank Acc. No. AF016509; Biswas and Russell, 1997. *J. Biol. Chem.* 272:15959-15966) is an isoform of the oxidative 17 beta-hydroxysteroid (17 beta-HSD6), and a human relative of rat retinol dehydrogenase (see *e.g.*, Chai, *et al.*, 1996. *Gene* 169:219-222). 17 beta-HSD6 is implicated in a wide range of disorders associated with retinol (*i.e.*, vitamin A) imbalance. Furthermore, 17 beta-HSD6 is implicated in disorders associated with androsterone deficiency.

(XII) Set* , GluT1* , CD24 and 115393

The present invention also discloses novel splice variants of Set and GluT1 (the proteins which are encoded are hereinafter designated as Set* and GluT1*, respectively), as well as novel open reading frames with the genes for CD24 and 115392 (the proteins which are encoded are

hereinafter designated as CD24* and 115392*, respectively), which interact with RB, as disclosed in this invention.

(a) CD24

The hybrid signal transduction protein CD24 (GenBank Acc. No. L33930; Kay, *et al.*, 1991. *J. Immunol.* 147:1412-1416) is a glycoprotein located on the surface of most B-lymphocytes. CD24 functions to inhibit differentiation and activation of antibody-producing cells; it is also implicated in the differentiation of granulocytes. It is also believed to transduce extracellular signals by release of second messengers, including calcium. See *e.g.*, Lund-Johansen, *et al.*, 1993. *Eur. J. Immunol.* 23:2782-2791. Furthermore, CD24 has been shown to interact with the members of the protein tyrosine kinase family (c-*fgr* and *lyn* products) and, thus, it may function to mediate signaling via phosphorylation cascades. See *e.g.*, Zarn, *et al.*, 1996. *Biochem. Biophys. Res. Commun.* 225:384-391. Interestingly, CD24 is also highly expressed on the surface of small cell lung carcinomas, thus tending to implicate this protein in tumorigenesis. See *e.g.*, Jackson, *et al.*, 1992. *Cancer Res.* 52:5264-5270.

(b) Set

The human *Set* gene (GenBank Acc. No. M93651; von Lindern, *et al.*, 1992. *Mol. Cell. Biol.* 12:3346-3355) encodes a putative oncogene-stimulating protein, Set. The *Set* gene has been localized to chromosomal region 9q34, which is associated with with various human leukemias . See *e.g.*, Adachi, *et al.*, 1994. *J. Biol. Chem.* 269:2258-2262. SET and a smaller fusion product (SET-CAN) is located to the nucleus, in specific association with the nuclear pore complex. SET-CAN is believed to regulate nucleocytoplasmic transport and cell cycle progression. See *e.g.*, Fornerod, *et al.*, 1996. *Oncogene* 13:1801-1808.

(c) GluT1

The human glutamate transporter protein (hereinafter designated GluT1; GenBank Acc. No. D26443; Nomura, *et al.*, 1994. *DNA Res.* 1:223-229), is an astroglial specific glutamate transporter. Alterations in astroglial glutamate transport have been implicated in the pathogenesis of hepatic encephalopathy and, possibly, in Alzheimer's Disease.

(d) clone 115392

The human mRNA clone 115392 (GenBank Accession No. L43578; Timms, *et al.*, 1996. *Genome Res.* 5:71-78) is an expressed sequence associated with Hunter Syndrome

(mucopolysaccharidosis type II), a fatal X-linked recessive disorder. However, the translation product for this mRNA has not been currently defined, and, thus, no function has yet been assigned.

SUMMARY OF THE INVENTION

In brief, the retinoblastoma protein (RB) has been demonstrated to form complexes, which heretofore have not been described. The present invention discloses herein compositions and methodologies for the production of protein complexes comprised of RB with proteins which interact with (*i.e.*, bind to) said RB protein. The proteins which have been demonstrated to form complexes with the RB protein will be designated hereinafter as "RB-IP" (for an RB-interacting protein) whereas a complex of the RB protein and an RB-IP will be hereinafter be designated as "RB•RB-IP".

More specifically, the present invention relates to complexes of RB, and derivatives, fragments and analogs thereof, with the following proteins: RN-tre, cellular apoptosis susceptibility protein CAS, gamma-interferon inducible protein IP-30, ribosomal protein L6, cytokine IK factor, lactate dehydrogenase B (LDH-B), Nlk1(Nek 2), cyclophilin A, Zap 3, oxidoreductase 17-beta-hydroxysteroid (17 beta-HSD6), CD24*, Set*, glutamate transporter GluT1* or clone 115392*, and their derivatives, analogs and fragments. The present invention discloses further the identification of the novel proteins, Set* and GluT1*, which are encoded by mRNA splice variants of the *Set* and *GluT1* genes, respectively. These aforementioned splice variants generated by the mRNAs encoding Set* and GluT1* are produced by the RNA splicing at splice sites other than the processing sites used to process the mRNAs encoding Set and GluT1, respectively. Additionally, the present invention discloses a novel protein encoded by an open reading frame (ORF) in the CD24 mRNA which is not part of, and does not overlap with, the nucleotide sequence coding for the known CD24 protein.

The present invention further discloses the nucleotide and amino acid sequences of Set* and GluT1* (the human Set* and GluT1* and homologs of other species), as well as derivatives (*e.g.*, fragments) and analogs thereof. The present invention also discloses the amino acid sequence of CD24* and 115392*, and derivatives (including fragments and analogs thereof). Nucleic acids able to hybridize to or complementary to the foregoing nucleotide sequences, such as the inverse complement (*i.e.*, a nucleic acid that has the complementary sequence running in reverse orientation to the strand so that the inverse complement would hybridize without mismatches to the nucleic acid strand) of the foregoing sequences, are also provided. In

particular, the present invention discloses nucleic acids which are hybridizable to or complementary to (*e.g.*, being the inverse complement) the portions of the nucleotide sequences encoding Set* and GluT1* which span the alternate splice junctions of the *Set** and *GluT1** mRNAs, respectively (*i.e.*, the point in the *Set** or *GluT1** nucleotide sequence at which the 5' and 3' splice sites were joined in processing the *Set** and *GluT1** mRNAs).

The present invention also discloses CD24*, Set*, GluT1*, and 115392* derivatives and analogs which possess functional biological activity (*i.e.*, they are capable of displaying one or more known functional biological activities of a wild-type CD24*, Set*, GluT1*, or 115392* protein). Such functional biological activities include, but are not limited to: (*i*) the ability to bind with, or compete for, interaction with RB; (*ii*) antigenicity (*i.e.*, the ability to bind or compete with CD24*, Set*, GluT1*, or 115392* for binding to an anti-CD24*, anti-Set*, anti-GluT1* or anti-115392* antibody, respectively) and (*iii*) immunogenicity (*i.e.*, the ability to generate an antibody that binds CD24*, Set*, GluT1*, or 115392*, respectively).

Methods of the production of the RB•RB-IP complexes and of CD24*, Set*, GluT1*, and 115392* proteins, and derivatives and analogs of the complexes and/or individual proteins by, for example, recombinant means will also be disclosed herein. Additionally, various pharmaceutical compositions relating to the RB•RB-IP complexes and of CD24*, Set*, GluT1*, and 115392* proteins, and derivatives and analogs of the complexes and/or individual proteins will also be disclosed by the present invention.

The present invention will further provide methodologies for the modulation (*i.e.*, inhibiting or enhancing) the activity of RB•RB-IP complexes, and CD24*, Set*, GluT1*, and 115392* proteins. The protein components of these aforementioned complexes have been implicated in various cellular functions, including, but not limited to: (*i*) physiological processes (*e.g.*, cell cycle control, cellular differentiation and apoptosis); (*ii*) response to viral infection; (*iii*) intracellular signal transduction; (*iv*) transcriptional control; and (*v*) pathophysiological processes (*e.g.*, hyperproliferative disorders including tumorigenesis and tumor spread, degenerative disorders including neurodegenerative diseases, virus infection).

Accordingly, the present invention discloses methodologies for the screening RB•RB-IP complexes, and CD24*, Set*, GluT1*, and 115392* proteins, as well as derivatives and analogs of the RB•RB-IP complexes, CD24*, Set*, Glut1*, and 115392* mRNAs, and CD24*, Set*, GluT1*, and 115392* proteins for their ability to alter cell functions, particularly those cell functions in which RB and/or an RB-IP has been implicated, as non-exclusively listed *supra*.

The present invention also discloses therapeutic and prophylactic, as well as diagnostic, prognostic, and screening methodologies and compositions which are based upon RB•RB-IP

complexes (as well as the nucleic acids encoding the individual proteins which participate in these aforementioned RB•RB-IP complexes) as well as CD24*, Set*, GluT1*, and 115392* proteins and nucleic acids. Therapeutic compounds of the present invention include, but are not limited to: (i) RB•RB-IP complexes and complexes where one or both members of the complex is a derivative or analog of RB or an RB-IP; CD24*, Set*, GluT1*, or 115392* proteins and derivatives, fragments and analogs thereof; (ii) antibodies to and nucleic acids encoding the foregoing; and (iii) antisense nucleic acids to the nucleotide sequences encoding the complex components and *CD24**, *Set**, *GluT1**, or *115392** antisense nucleic acids. Diagnostic, prognostic and screening kits will also be provided.

Animal models and methods of screening for modulatory agents (*i.e.*, agonists, antagonists and inhibitors) of the activity of RB•RB-IP complexes, and CD24*, Set*, GluT1*, and 115392* proteins, will also be disclosed by the present invention.

Methodologies for the identification of molecules which possess the ability to inhibit, or alternatively, which increase formation of RB•RB-IP complexes will also be provided.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1: illustrates the *CD24* nucleotide acid sequence (GenBank Accession No. L33930; [SEQ ID NO:1]). The initiation methionine codon ATG of CD24 is boxed at the start of the CD24 coding sequence, which is underlined and denoted by callout "**A**". The translational stop codon TAA used for CD24 is boxed at the end of the underlined sequence. The prey sequence identified in the assay described in the Specific Examples Section, *infra*, begins at nucleotide 689, denoted by the arrow and callout "**B**". In addition, potential open reading frames (ORFs) included in the identified prey sequence are denoted by callouts "**C**", "**D**", and "**E**". It should be noted, however, that only the open reading frame (ORF) of callouts "**A**" and "**C**" represent potential *in vivo* protein products, as determined in the Specific Examples Section, *infra*.

Figure 2: illustrates the CD24 nucleotide and amino acid sequences (Frame A; [SEQ ID NOS:2 and 3, respectively), and the CD24* nucleotide and amino acid sequences (Frame B; [SEQ ID NOS:4 and 5, respectively). Nucleotide residue numbers retain the original schema of the CD24 sequence displayed in Figure 1, while amino acids are numbered consecutively for each individual protein.

Figure 3: illustrates the nucleotide sequence of *Set* (GenBank Accession No. M93651; [SEQ ID NO:6]). The known Set protein-encoding sequence is underlined, flanked by the initiation codon ATG (denoted by callout "**A**") and the termination codon TAA (denoted by callout "**C**"). The prey sequence identified in the Specific Examples Section *infra,* begins at nucleotide 2076, donated by the white arrow and callout "**F**", and falls outside of the known coding sequence for Set. Nucleotides which are identical to the known consensus sequence for 5'- splice sites are shown in bold, and the last base of the exon (exon 1) is denoted by an arrow and callout "**B**". A 3'- splice site, with nucleotides identical to the known consensus sequence for 3'- splice sites, is shown in bold, and the first base of the exon (exon 2) is denoted by an arrow and callout "**E**". The stop codon TAA for Set* is denoted by callout "**G**". The branch point consensus sequence for this exon, with bases matching the canonical branch point consensus bases shown in bold, is denoted by callout "**D**".

Figure 4: illustrates the Set* nucleotide sequence [SEQ ID NO:7] and associated amino acid sequence [SEQ ID NO:8]. The amino acid residue at which the sequence deviates from that of Set because of alternate splicing is denoted by arrow "**A**". The prey sequence identified in the assay described in the Specific Examples Section, *infra*, begins at nucleotide 2076 of the *Set* sequence (Figure 3), and is denoted by arrow "**B**".

Figure 5: illustrates the nucleotide sequence of *GluT1* (GenBank Accession No. D26443; [SEQ ID NO:9]). The sequence of the known GluT1 coding sequence is underlined and flanked by the boxed initiator methionine ATG codon (denoted by callout "**A**") and the boxed translational stop codon TAG (denoted by callout "**C**"). The prey sequence identified in the Specific Examples Section *infra,* begins at nucleotide 3408, denoted by arrow and callout "**F**", and falls outside of the known coding sequence for GluT1. A 5'- splice site, with nucleotides identical to the known consensus sequence for 5'- splice sites shown in bold, and the last nucleotide of the exon (exon 1), is denoted by an arrow and callout "**B**". A compatible 3'- splice site, with nucleotides identical to the known consensus sequence for 3'- splice sites shown in bold, and the first base of the exon (exon 2) is denoted by an arrow and callout "**E**". The stop codon TAA for GluT1* is denoted by callout "**G**". A branch point consensus sequence for this exon, with bases matching the consensus bases shown in bold, and is denoted by callout "**D**".

Figure 6: illustrates the GluT1* nucleotide sequence [SEQ ID NO:10] and associated amino acid sequence [SEQ ID NO:11]. The amino acid residue at which the sequence deviates from

that of GluT1 because of alternate splicing is denoted by arrow "**A**". The prey sequence identified in the assay described in the Specific Examples Section, *infra*, begins at nucleotide 1789 of the original GluT1 sequence (Figure 5), denoted by arrow "**B**".

Figure 7: illustrates the 115392* nucleotide sequence (GenBank Accession No. L43578; [SEQ ID NO:12]) and associated amino acid sequence [SEQ ID NO:13]. The prey sequences identified in the assay described in the Specific Examples Section, *infra*, begins at nucleotide 343 (amino acid residue 33), and is denoted by arrow "**A**". The open reading frame (ORF) was determined as described *infra,* in Sections (II) and the Specific Example Section.

Figure 8: illustrates the specificity of RB interactions. Shown is the matrix of results of the yeast two hybrid system assays. The results of assays using the bait proteins RB and B1 are indicated to the left of the rows, and the prey proteins RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, 115392*, RbAP46 and P1 are indicated above the columns. A positive interaction between the indicated bait and prey proteins is indicated as "+" in the box forming the intersection between the particular bait and prey proteins; a lack of interaction is designated by an empty box. Boxes labeled A through P indicate the results of matings and growth of yeast expressing RB and RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, 115392*, respectively. The box labeled Q indicates the mating and growth of yeast expressing bait protein RB and protein Rb AP46, recapitulating a previously known interaction. See *e.g.*, Hwang, *et al.*, 1991. *Nature* 350:160-162.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is based upon the identification of proteins which possess the ability to interact with RB (hereinafter referred to as "RB-IPs") using an improved, modified form of the yeast two hybrid system. In the present invention, RN-tre, cellular apoptosis susceptibility protein CAS, gamma-interferon inducible protein IP-30, ribosomal protein L6, cytokine IK factor, lactate dehydrogenase B (LDH-B), Nlk1(Nek 2), cyclophilin A, Zap 3, oxidoreductase 17-beta-hydroxysteroid (17 beta-HSD6), CD24*, Set*, glutamate transporter GluT1* or clone 115392* (wherein "*" indicates splice variants of Set and GluT1, and novel open reading frames (ORFs) of CD24 and 115392* as determined in the Specific Examples Section, *infra*, respectively) were found to form complexes under physiological conditions with

RB. These aforementioned complexes of RB with an RB-IP are hereinafter designated "RB•RB-IP" complexes. These RB•RB-IP complexes, by virtue of this interaction, are implicated in the modulation of various functional activities of RB, and of its binding partners. These functional activities include, but are not limited to: (*i*) physiological processes (*e.g.*, cell cycle control, cellular differentiation and apoptosis); (*ii*) response to viral infection; (*iii*) intracellular signal transduction; (*iv*) transcriptional control; and (*v*) pathophysiological processes (*e.g.*, hyperproliferative disorders including tumorigenesis and tumor spread, degenerative disorders including neurodegenerative diseases, virus infection).

The present invention discloses methodologies for the screening of proteins which interact with (*e.g.*, bind to) RB. The present invention further relates to RB complexes (RB•RB-IPs), in particular RB complexed with one of the following proteins: RN-tre, cellular apoptosis susceptibility protein CAS, gamma-interferon inducible protein IP-30, ribosomal protein L6, cytokine IK factor, lactate dehydrogenase B (LDH-B), Nlk1(Nek 2), cyclophilin A, Zap 3, oxidoreductase 17-beta-hydroxysteroid (17 beta-HSD6), CD24*, Set*, glutamate transporter GluT1* or clone 115392*. The present invention further discloses complexes of RB or derivatives, analogs and fragments of RB, with RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, and 115392*, or derivatives, analogs and fragments thereof. In a preferred embodiment such complexes bind an anti-RB•RB-IP complex antibody. In a specific embodiment of the present invention, complexes of human RB with human proteins are disclosed.

The present invention also discloses methodologies for the production and/or isolation of RB•RB-IP complexes. In a specific embodiment, the present invention provides methods of using recombinant DNA techniques to express both RB and its binding partner (RB-IP), or fragments, derivatives or homologs of one or both members of the complex, either where both binding partners are under the control of one heterologous promoter (*i.e.*, a promoter not naturally associated with the native gene encoding the particular complex component) or where each is under the control of a separate heterologous promoter.

As will be described *infra*, the present invention discloses novel proteins encoded by novel splice variants of the *Set* and *GluT1* genes, which encode proteins designated Set* and GluT1*, that also possess the ability to interact with RB. The nucleotide and associated amino acid sequences of Set* and GluT1* are depicted in Figures 3-4 and 5-6, respectively [SEQ ID NOS:6-8 and 9-11, respectively]. The present invention also discloses a protein which interacts with RB, and is encoded by an ORF found within the *CD24* gene which does not overlap with the nucleotide sequence encoding the CD24 protein. The nucleotide and associated amino acid

sequences of CD24* are depicted in Figure 2 [SEQ ID NOS:4 and 5, respectively]. Additionally, the present invention discloses a protein encoded by a nucleotide sequence within the EST sequence 115392 (the protein is hereinafter designated 115392*) which interacts with RB. The nucleotide sequence of EST sequence 115393* and its associated amino acid sequence are depicted in Figure 6 [SEQ ID NOS:12 and 13, respectively].

In another aspect, the present invention discloses the nucleotide sequences of *Set** and *GluT1** and their encoded proteins, as well as the mRNA nucleotide sequence (including the DNA sequence corresponding to the mRNA sequence) and associated amino acid sequence of CD24*. The present invention further discloses CD24*, Set*, GluT1*, and 115392* proteins, and derivatives (including but not limited to fragments) and homologs and paralogs thereof, as well as nucleic acids encoding the CD24*, Set*, GluT1*, and 115392* proteins, derivatives, fragments and homologs. The present invention also provides CD24* protein, and Set*, GluT1*, and 115392* proteins and genes encoding these proteins, of many different species, preferably vertebrates and more preferably mammals. In the most preferred embodiment, the CD24* protein, and the Set*, GluT1*, and 115392* proteins, and genes are of human origin. The production of the aforementioned proteins and derivatives are also disclosed (*e.g.*, by recombinant methodologies).

The present invention further relates to CD24*, Set*, GluT1*, and 115392* derivatives and analogs that possess functional biological activity (*i.e.*, are capable of displaying one or more known functional activities associated with a full-length, wild-type CD24*, Set*, GluT1*, or 115392*) Such functional biological include, but are not limited to: (*i*) the ability to form a complex with RB; (*ii*) antigenicity (*i.e.*, the ability to bind, or compete with CD24*, Set*, GluT1*, or 115392* for binding, to an anti-CD24*, anti-Set*, anti-GluT1*, or anti-115392* antibody, respectively); (*iii*) immunogenicity (*i.e.*, the ability to generate an antibody that binds to CD24*, Set*, GluT1*, or 115392*, respectively) and the like.

Methods of diagnosis, prognosis, and screening for diseases and disorders associated with aberrant levels of RB•RB-IP complexes, or of CD24*, Set*, GluT1*, or 115392* are also disclosed in the present invention. Furthermore, the present invention provides methods of treating or preventing diseases or disorders associated with aberrant levels of RB•RB-IP complexes or CD24*, Set*, GluT1*, or 115392*, or aberrant levels of the activity of one or more of the components of the complex by administration of the RB•RB-IP complexes, CD24*, Set*, GluT1*, or 115392*, or modulators of RB•RB-IP complex formation or activity, mutants of RB or the RB-IP which increase or decrease binding affinity, small molecule inhibitors/enhancers of complex formation, antibodies that either stabilize or neutralize the complex, and the like.

Methods of assaying RB•RB-IP complexes, CD24*, Set*, GluT1*, or 115392*, for activity as therapeutics or diagnostics, as well as methods of screening for RB•RB-IP complex, CD24*, Set*, GluT1*, or 115392* modulators (*i.e.*, inhibitors, agonists and antagonists) are also disclosed herein.

For clarity of disclosure, and not by way of limitation, the following detailed description of the present invention will be divided into the various sub-sections *infra*.

(I) RB•IP Complexes and CD24*, Set*, GluT1*, and 115392* Proteins, Derivatives and Analogs

The present invention discloses RB•RB-IP complexes and, in particular aspects, complexes of RB and RN-tre, RB and CAS, RB and IP-30, RB and L6, RB and IK, RB and LDH-B, RB and Nlk1(Nek 2), RB and cyclophilin A, RB and Zap 3, RB and 17 beta-HSD6, RB and CD24*, RB and Set*, RB and GluT1*, and RB and 115392*. In a preferred embodiment of the present invention, the RB•RB-IP complexes are complexes of human proteins. The present invention also discloses complexes of derivatives (including fragments) and analogs of RB with an RB-IP, complexes of RB with derivatives (including fragments) and analogs of an RB-IP, and complexes of derivatives (including fragments) and analogs of RB and derivatives (including fragments) and analogs of an RB-IP. As utilized herein, fragments, derivatives or analogs of an RB•RB-IP complex include complexes where one or both members of the complex are fragments, derivatives or analogs of the wild-type RB or RB-IP protein. Preferably, the RB•RB-IP complexes in which one or both members of the complex are a fragment, derivative or analog of the wild type protein are functional biologically active RB•RB-IP complexes. In specific embodiments, the native proteins, derivatives or analogs of RB and/or the RB-IP are derived from animals or plant sources. "Functional biologically active RB•RB-IP complexes" as used herein, refers to those molecules displaying one or more of the known functional biological attributes of a complex of full length, wild-type RB with a full length, wild-type RB-IP (*e.g.*, RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, or 115392*), including, but not limited to (*i*) physiological processes (*e.g.*, cell cycle control, cellular differentiation and apoptosis); (*ii*) response to viral infection; (*iii*) intracellular signal transduction; (*iv*) transcriptional control; and (*v*) pathophysiological processes (*e.g.*, hyperproliferative disorders including tumorigenesis and tumor spread, degenerative disorders including neurodegenerative diseases, virus infection).

Accordingly, the present invention provides methods for the screening of RB•RB-IP complexes, particularly complexes of RB with RN-tre, CAS, IP-30, L6, IK, LDH-B,

Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, or 115392*, and the CD24*, Set*, GluT1*, and 115392* proteins, as well as derivatives and analogs of the RB•RB-IP complexes, and CD24*, Set*, GluT1*, and 115392* proteins, for the ability to alter cell functions, particularly those cell functions in which RB and/or an RB-IP has been implicated, such as but not limited to: (*i*) physiological processes (*e.g.*, cell cycle control, cellular differentiation and apoptosis); (*ii*) response to viral infection; (*iii*) intracellular signal transduction; (*iv*) transcriptional control; and (*v*) pathophysiological processes (*e.g.*, hyperproliferative disorders including tumorigenesis and tumor spread, degenerative disorders including neurodegenerative diseases, virus infection)., binding to an anti-RB•RB-IP complex antibody; and various other activities as known within the art. For example, such derivatives or analogs which possess the desired immunogenicity or antigenicity may be used in immunoassays, for immunization, for inhibition of RB•RB-IP complex activity, and the like. Derivatives or analogs which retain or enhance, or alternatively lack or inhibit, a specific property of interest (*e.g.*, participation in an RB•RB-IP complex) may be utilized as inducers, or inhibitors, respectively, of such a property and its physiological correlates. An other specific embodiment of the present invention relates to an RB•RB-IP complex of a fragment of an RB and/or a fragment of RB-IP which may be bound by an anti-RB and/or anti-RB-IP antibody or antibody specific for an RB•RB-IP complex when such a fragment is included within a given RB•RB-IP complex.

Fragments and other derivatives or analogs of RB•RB-IP complexes can be tested for the desired activity by procedures well-known within the art, including, but not limited to those assays described in Section (VI), *infra*.

In a specific embodiment, the present invention provides RB•RB-IP complexes comprising fragments of one or both members of the complex. In a preferred embodiment, these fragments consist of, but are not limited to, fragments of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, or 115392*, identified as interacting with RB in the improved, modified yeast two hybrid assay. These aforementioned fragments include, but are not limited to: amino acid residues 758-827 of RN-tre; amino acid residues 714-971 of CAS; amino acid residues 142-303 and 146-303 of IP-30; amino acid residues 64-288 of L6; amino acid residues 1-162 of IK; amino acid residues 265-334 of LDH-B; amino acid residues 319-445 of Nlk1(Nek 2); amino acid residues 1-165 of cyclophilin A; amino acid residues 296-331 of Zap 3; amino acid residues 1-317 of 17 beta-HSD6; amino acids 1-64 of CD24* as depicted in Figure 2 [SEQ ID NO:6]; amino acid residues 105-116 of Set* as depicted in Figure 4 [SEQ ID NO:8]; amino acid residues 297-316 of GluT1* as depicted in Figure 6

[SEQ ID NO:11] and amino acid residues 33-158 of 115392* as depicted in Figure 7 [SEQ. ID. NO. 13]. Additionally, fragments, or proteins comprising fragments, which lack some or all of the aforementioned amino acid residues of either member of the RB•RB-IP complex assay, are also provided. Moreover, nucleic acids which encode the foregoing are also disclosed herein.

The present invention further relates to CD24*, Set*, GluT1*, and 115392* proteins as well as derivatives and homologs and paralogs of CD24*, Set*, GluT1*, and 115392* proteins. In one embodiment of the present invention, human CD24*, Set*, GluT1*, and 115392* genes and proteins are disclosed. In specific embodiments, the native proteins, fragments, derivatives or analogs of CD24*, Set*, GluT1*, or 115392* are derived from animals (*e.g.*, mouse, rat, pig, cow, dog, monkey, human, fly, or frog) or from plants. In other specific embodiments, the fragment, derivative or analog is possesses functional biological activity (*i.e.*, is capable of exhibiting one or more functional activities associated with full-length, wild-type CD24*, Set*, GluT1*, or 115392*, for example, the ability to bind RB, immunogenicity or antigenicity).

The nucleotide sequences encoding human RB, RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24, Set, GluT1, and clone 115392 are known (GenBank Accession No. M28419; GenBank Accession No. D13644; GenBank Accession No. U33286; GenBank Accession No. J03909; GenBank Accession No. X69391; GenBank Accession No. S74221; GenBank Accession No. Y00711; GenBank Accession No. U11050; GenBank Accession No. Y00052; GenBank Accession No. L40403; GenBank Accession No. AF016509; GenBank Accession No. L33930; GenBank Accession No. M93651; GenBank Accession No. D26443; and GenBank Bank Accession No. L43578, respectively). The nucleotide sequences encoding CD24, Set, GluT1, and clone 115392 are provided in Figures 1, 3, 5, and 7 (SEQ ID NOS: 1, 6, 9, and 12), respectively. It should be noted that the nucleic acids encoding RB, RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, or clone 115392* may be readily obtained by any method known within the art (*e.g.*, by PCR amplification using synthetic primers hybridizable to the 3'- and 5'-termini of the sequence and/or by cloning from a cDNA or genomic library using an oligonucleotide specific for the gene sequence. These aforementioned methodologies are fully described in Section (II), *infra*.

Homologs (*i.e.*, those nucleic acids encoding RB, RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, or 115392* of species other than human) or other related sequences (*e.g.*, paralogs) may be obtained by low, moderate or high stringency hybridization with all or a portion of the particular human sequence as a probe

using methods well-known within the art for nucleic acid hybridization and cloning (*e.g.*, as described in Section (II) for *CD24*, Set*, GluT1*, and 115392** sequences).

Similarly, the RB, RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, or 115392* proteins either alone or in a complex, may be obtained by methods well-known within the art for protein purification and recombinant protein expression. For example, for recombinant expression of one or more of the proteins, the nucleic acid containing all or a portion of the nucleotide sequence encoding the protein may be inserted into an appropriate expression vector (*i.e.*, a vector that contains the necessary elements for the transcription and translation of the inserted protein coding sequence). Additionally, the necessary transcriptional and translational signals may also be supplied by the native promoter for RB or any RB-IP genes, and/or their flanking regions.

A variety of host-vector systems may be utilized to express the protein coding sequence. These host vector systems include, but are not limited to, mammalian cell systems infected with virus (*e.g.*, vaccinia virus, adenovirus, etc.); insect cell systems infected with virus (*e.g.*, baculovirus); microorganisms such as yeast containing yeast vectors; or bacteria transformed with bacteriophage, DNA, plasmid DNA, or cosmid DNA. The expression elements of vectors vary in their overall strengths and specificities. Depending on the host-vector system utilized, any one of a number of suitable transcription and translation elements may be utilized.

In a preferred embodiment of the present invention, the RB•RB-IP complexes are obtained by expressing the entire RB sequence and an RB-IP coding sequence in the same cell, either under the control of the same promoter or two separate promoters. In yet another embodiment, a derivative, fragment or homolog of RB and/or a derivative, fragment or homolog of an RB-IP are recombinantly expressed. Preferably the derivative, fragment or homolog of RB and/or the RB-IP protein forms a complex with a binding partner identified by a binding assay, such as the modified yeast two hybrid system, and more preferably forms a complex that binds to an anti-RB•RB-IP complex antibody.

Any of the methods described for the insertion of DNA fragments into a vector may be used to construct expression vectors containing a chimeric gene consisting of appropriate transcriptional/translational control signals and protein coding sequences. These methods may include *in vitro* recombinant DNA and synthetic techniques and *in vivo* recombinants (genetic recombination). Expression of nucleotide sequences encoding RB and an RB-IP, or derivatives, fragments or homologs thereof, may also be regulated by a second nucleotide sequence so that the gene(s) or fragment(s) thereof are expressed in a host transformed with the recombinant DNA molecule(s). For example, expression of the proteins may be controlled by any

promoter/enhancer known within the art. In a specific embodiment of the present invention, the promoter which is utilized is not native (*i.e.*, is not the wild-type promoter) to the genes for RB or the specific RB-IP. Promoters which may be utilized include, but are not limited to: (*i*) the SV40 early promoter (see *e.g.*, Bernoist & Chambon, 1981. *Nature* 290:304-310); (*ii*) the promoter contained in the 3'- long terminal repeat of Rous sarcoma virus (see *e.g.*, Yamamoto, *et al.*, 1980. *Cell* 22:787-797); (*iii*) the Herpesvirus thymidine kinase promoter (see *e.g.*, Wagner, *et al.*, 1981. *Proc. Natl. Acad. Sci. USA* 78:1441-1445); (*iv*) the regulatory sequences of the metallothionein gene (see *e.g.*, Brinster, *et al.*, 1982. *Nature* 296:39-42); (*v*) prokaryotic expression vectors (see *e.g.*, Villa-Kamaroff, *et al.*, 1978. *Proc. Natl. Acad. Sci. USA* 75:3727-3731) or (*vi*) the *tac* promoter (see *e.g.*, DeBoer, *et al.*, 1983. *Proc. Natl. Acad. Sci. USA* 80:21-25).

In addition, animal transcriptional control regions which exhibit tissue specificity and have been utilized in transgenic animals may also be utilized. These transcriptional control regions include, but are not limited to: (*i*) the elastase I gene control region which is active in pancreatic acinar cells (see *e.g.*, Swift, *et al.*, 1984. *Cell* 38:639-646; (*ii*) the insulin gene control region which is active in pancreatic β-cells (see *e.g.*, Hanahan, *et al.*, 1985. *Nature* 315:115-122); (*iii*) the immunoglobulin gene control region which is active in lymphoid cells (see *e.g.*, Alexander, *et al.*, 1987. *Mol. Cell Biol.* 7:1436-1444); (*iv*) the mouse mammary tumor virus control region which is active in testicular, breast, lymphoid and mast cells (see *e.g.*, Leder, *et al.*, 1986. *Cell* 45:485-495); (*v*) the α-fetoprotein gene control region which is active in liver (see *e.g.*, Krumlauf, *et al.*, 1985. *Mol. Cell. Biol.* 5:1639-1648); (*vi*) the β-globin gene control region which is active in myeloid cells (see *e.g.*, Kollias, *et al.*, 1986. *Cell* 46:89-94); (*vii*) the myosin light chain-2 gene control region which is active in skeletal muscle (see *e.g.*, Sani, 1985. *Nature* 314:283-286) and (vii) gonadotrophic releasing hormone gene control region which is active in gonadotrophs of the hypothalamus (see *e.g.*, Mason, *et al.*, 1986. *Science* 234:1372-1378).

In a specific embodiment of the present invention, a vector is used which comprises: (*i*) a promoter operably-linked to nucleotide sequences encoding RB and/or an RB-IP (*e.g.*, RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, or 115392*), or a fragment, derivative or homolog, thereof; (*ii*) one or more origins of replication and, optionally; (*iii*) one or more selectable markers (*e.g.*, an antibiotic resistance gene). In a preferred embodiment, a vector is utilized which comprises a promoter operably-linked to nucleotide sequences encoding both RB and an RB-IP, one or more origins of replication, and optionally, one or more selectable markers.

In another specific embodiment of the present invention, an expression vector containing the coding sequences, or portions thereof, of RB and an RB-IP (*e.g.*, RN-tre, CAS, IP-30, L6, IK,

LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, or 115392*), either together or separately, is constructed, by way of example but not of limitation, by initially sub-cloning the gene sequences into the EcoRI restriction site of each of the three pGEX vectors (glutathione S-transferase expression vectors; Promega Corp.; Madison, WI.; see also Johnson, 1988. *Gene* 7:31-40). This aforementioned cloning system allows for the expression of selected gene products in the correct reading frame.

Expression vectors containing the sequences of interest may be identified by three general methodologies: (*i*) by nucleic acid hybridization; (*ii*) by the presence or absence of "marker" gene function and/or (c) by the expression of the inserted sequences. In the first methodology, RB, RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, and 115392*, or other RB-IP sequences may be detected by nucleic acid hybridization to probes comprising sequences homologous and complementary to the inserted sequences. In the second methodology, the recombinant vector/host system can be identified and selected based upon the presence or absence of certain "marker" functions (*e.g.*, binding to an anti-RB, anti-RB-IP, or anti-RB•RB-IP complex antibody, resistance to antibiotics, occlusion body formation in baculovirus, and the like) caused by insertion of the sequences of interest in the vector. For example, if an *RB* or *RB-IP* gene, or portion thereof, is inserted within the marker gene sequence of the vector, recombinants containing the *RB* or *RB-IP* fragment will be identified by the absence of the marker gene function. In the third approach, recombinant expression vectors can be identified by assaying for RB, RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, and 115392*, or other RB-IP products expressed by the recombinant vector. Such assays can be based, for example, on the physical or functional properties of the interacting species in *in vitro* assay systems (*e.g.*, formation of a RB•RB-IP complex or immunoreactivity to antibodies specific for the protein).

Subsequently, once the recombinant RB, RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, and 115392*, or other RB-IP molecules are identified and the complexes or individual proteins are isolated, various methods well-known within the art may be used to propagate them. Moreover, when a suitable host system and growth conditions have been established, recombinant expression vectors may then be propagated and amplified in quantity. As previously described, the expression vectors or derivatives which can be used include, but are not limited to: human or animal viruses; insect viruses; yeast vectors; bacteriophage vectors and plasmid and cosmid vectors.

Additionally, a host cell strain may be chosen that modulates the expression of the inserted sequences, or modifies or processes the expressed proteins in the specific fashion

desired. Expression from certain promoters can be elevated in the presence of certain inducers; thus expression of the genetically-engineered *RB* and/or *RB-IP* may be controlled. Furthermore, different host cells have characteristic and specific mechanisms for the translational and post-translational processing and modification (*e.g.*, glycosylation, phosphorylation, etc.) of proteins. Appropriate cell lines or host systems can be chosen to ensure the desired modification and processing of the foreign protein is achieved. For example, expression in a bacterial system can be used to produce an non-glycosylated core protein, while expression in mammalian cells ensures "native" glycosylation of a heterologous protein. Furthermore, different vector/host expression systems may effect processing reactions to different extents.

In other specific embodiments, the RB and/or RB-IPs, or fragments, homologs or derivatives thereof, may be expressed as fusion or chimeric protein products comprising the protein, fragment, homolog, or derivative joined via a peptide bond to a heterologous protein sequence of a different protein. Such chimeric products can be made by ligating the appropriate nucleic acids encoding the desired amino acids to each other in the proper coding frame by methods known in the art, and expressing the chimeric products in a suitable host by methods commonly known in the art. Alternatively, such a chimeric product can be made by protein synthetic techniques (*e.g.*, by use of a peptide synthesizer). Chimeric genes comprising portions of RB and/or an RB-IP, or CD24*, Set*, GluT1*, or 115392*, fused to any heterologous protein-encoding sequences may be constructed. A specific embodiment of the present invention discloses a chimeric protein comprising a fragment of RB and/or an RB-IP, or CD24*, Set*, GluT1*, or 115392*, of at least six amino acid residues in length.

In another specific embodiment of the present invention, fusion proteins are disclosed which possess the interacting domains of the RB protein and an RB-IP (*e.g.*, RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, and 115392*) and/or, optionally, a hetero-functional reagent (*e.g.*, a peptide linker between the two domains) where such a reagent promotes the interaction of the respective RB and RB-IP binding domains. These fusion proteins may be particularly useful where the stability of the interaction is desirable due to the formation of the complex as an intra-molecular reaction (*e.g.*, in the production of antibodies specific to the RB•RB-IP complex).

More specifically, RB and/or derivatives of an RB-IP, or CD24*, Set*, GluT1*, or 115392* may be generated by altering their respective sequences by substitutions, additions and/or deletions which provide for functionally-equivalent molecules. Due to the degeneracy of nucleotide coding sequences, other DNA sequences that encode substantially the same amino acid sequence as an *RB* or *RB-IP* gene can be used in the practice of the present invention. These

include, but are not limited to nucleotide sequences comprising all or portions of *RB, RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1**, and *115392** genes that are altered by the substitution of different codons that encode the same amino acid residue within the sequence, thus producing a "silent" alteration.

Similarly, RB and RB-IP derivatives of the present invention include, but are not limited to, those containing, as a primary amino acid sequence, all or part of the amino acid sequence of RB or an RB-IP, including altered sequences in which functionally equivalent amino acid residues are substituted for residues within the sequence resulting in a silent change. For example, one or more amino acid residues within the sequence can be substituted by another amino acid of a similar polarity which acts as a functional equivalent, resulting in a silent alteration. Substitutes for an amino acid within the sequence may be selected from other members of the class to which the amino acid belongs.

In a specific embodiment of the present invention, the nucleic acids encoding proteins, and proteins consisting of, or comprising a fragment of RB or an RB-IP comprised of at least 6 (contiguous) amino acid residues of RB or an RB-IP, are disclosed. In other specific embodiments, the aforementioned amino acid fragments are comprised of at least 10, 20, 30, 40, 50, 100 or 200 amino acids of RB or an RB-IP. Derivatives or analogs of RB and RB-IPs, include, but are not limited to, molecules comprising regions that are substantially homologous to RB or RB-IPs, in various embodiments, by at least 30%, 40%, 50%, 60%, 70%, 80%, 90% or 95% identity over an amino acid sequence of identical size or when compared to an aligned sequence in which the alignment is done by a computer homology program known within the art, or whose encoding nucleic acid is capable of hybridizing to the complement (*e.g.*, the inverse complement) of a sequence encoding RB or an RB-IP under stringent, moderately stringent, or non-stringent conditions, as described *supra,* Section (I).

The RB, RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, and 115392* derivatives and analogs of the present invention may be produced by various methods known within the art. The manipulations which result in their production can occur at the gene or protein level. For example, the cloned *RB* and *RB-IP* gene sequences can be modified by any of numerous strategies known in the art. See *e.g.,* Sambrook, *et al.,* 1989. *Molecular Cloning, A Laboratory Manual, 2d Ed.,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. For example, the sequences of interest may be cleaved at appropriate sites with restriction endonuclease(s), followed by further enzymatic modification if desired, isolated, and ligated *in vitro.* It should be noted, however, that in the production of the gene encoding a derivative or analog of RB or an RB-IP, care should be taken

to ensure that the modified gene retains the original translational reading frame, uninterrupted by translational stop signals, in the gene region where the desired activity is encoded.

Additionally, the RB and/or RB-IP-encoding nucleotide sequence can be mutated *in vitro* or *in vivo*, to create and/or destroy translation, initiation, and/or termination sequences, or to create variations in coding regions and/or form new restriction endonuclease sites or destroy pre-existing ones, to facilitate further *in vitro* modification. Any technique for mutagenesis known in the art can be used, including but not limited to, chemical mutagenesis and *in vitro* site-directed mutagenesis (see *e.g.,* Hutchinson, *et al.,* 1978. *J. Biol. Chem.* 253:6551-6558), the use of TAB™ linkers (Pharmacia, Upsala, Sweden), and the like.

Subsequently, once a recombinant cell expressing RB and/or an RB-IP protein, or fragment or derivative thereof, is identified, the individual gene product or complex can be isolated and analyzed. This is achieved by assays based on the physical and/or functional properties of the protein or complex, including, but not limited to, radioactive labeling of the product followed by analysis by gel electrophoresis, immunoassay, cross-linking to marker-labeled product, and the like.

The RB•RB-IP complexes, and CD24*, Set*, GluT1*, and 115392* proteins, may be isolated and purified by standard methodologies known within the art (either from natural sources or recombinant host cells expressing the complexes or proteins) including, but not limited to: (*i*) column chromatography (*e.g.,* ion exchange, affinity, gel exclusion, reversed-phase high pressure, fast protein liquid, etc.); (*ii*) differential centrifugation; (*iii*) differential solubility or (*iv*) by any other standard technique used for the purification of proteins. Functional properties may be evaluated using any suitable assay known in the art.

Alternatively, once an RB-IP or its derivative is identified, the amino acid sequence of the protein can be deduced from the nucleotide sequence of the chimeric gene from which it was encoded. As a result, the protein (or its derivative) may be synthesized by standard chemical methods known within the art. See, *e.g.,* Hunkapiller, *et al.,* 1984. *Nature* 310:105-111.

In a specific embodiment of the present invention, such RB•RB-IP complexes, as well as CD24*, Set*, GluT1*, and 115392* proteins, whether produced by recombinant DNA techniques, chemical synthesis methods, or by purification from native sources, include, but are not limited to those containing, as a primary amino acid sequence, all or part of the amino acid sequences, as well as fragments and other analogs and derivatives thereof, including proteins homologous thereto.

Manipulations of RB and/or RB-IP sequences may be made at the protein level. Included within the scope of the present invention are derivatives or complexes of RB or RB-IP fragments,

derivatives or analogs of CD24*, Set*, GluT1*, and 115392*, and RB, RB-IP, CD24*, Set*, GluT1*, and 115392* fragments, derivatives and analogs that are differentially modified during or after translation (*e.g.*, by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, and the like). In addition, any of numerous chemical modifications may also be performed by well-known techniques, including but not limited to, specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, $NaBH_4$, acetylation, formylation, oxidation, reduction, synthesis in the presence of tunicamycin, and the like.

In specific embodiments of the present invention, the RB and/or RB-IP sequences are modified to include a fluorescent label. In another specific embodiment, the RB and/or the RB-IP are modified so as to possess a heterofunctional reagent which may be utilized to cross-link the protein to other members of the complex or to other RP-IPs. In addition, analogs and derivatives of RB and/or an RB-IP, or analogs and derivatives of CD24*, Set*, GluT1*, or 115392*, can be chemically synthesized by use of a peptide synthesizer. Furthermore, if desired, non-classical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the RB and/or an RB-IP.

In cases where natural products are suspected of being mutant or are isolated from new species, the amino acid sequence of RB, or an RB-IP isolated from the natural source, as well as those expressed *in vitro*, or from synthesized expression vectors *in vivo* or *in vitro*, can be determined from analysis of the DNA sequence, or alternatively, by direct sequencing of the isolated protein. Such analysis may be performed by manual sequencing or through use of an automated amino acid sequenator. The RB•RB-IP complexes, or CD24*, Set*, GluT1*, or 115392* proteins, may also be analyzed by hydrophilicity analysis. See *e.g.*, Hopp and Woods, 1981. *Proc. Natl. Acad. Sci. USA* 78: 3824-3828. Secondary structural analysis may also be performed to facilitate the identification of specific regions of the RB and/or an RB-IP that assume specific structural motifs. See *e.g.*, Chou and Fasman, 1974. *Biochemistry* 13: 222-23. Computer software programs (available within the art) may also be utilized in various analyses including, but not limited to: manipulation, translation, secondary structure prediction, hydrophilicity and hydrophobicity profiles, open reading frame prediction and plotting, and determination of sequence homologies. Other methods of structural analysis including, but not limited to: X-ray crystallography (see *e.g.*, Engstrom, 1974. *Biochem. Exp. Biol.* 11:7-13); mass spectroscopy and gas chromatography (see *e.g.*, *Methods in Protein Science*, 1997. J. Wiley and Sons, New York) and computer modeling (see *e.g.*, Fletterick and Zoller, eds., 1986. Computer

Graphics and Molecular Modeling, In: *Current Communications in Molecular Biology*, Cold Spring Harbor Laboratory, Cold Spring Harbor Press, New York) may also be employed.

(II) Identification and Isolation of *CD24**, *Set**, *GluT1**, AND *115392** Genes

The present invention discloses the nucleotide sequences encoding CD24*, Set*, GluT1*, and 115392*. In specific embodiments, the *CD24**, *Set**, *GluT1**, and *115392** nucleic acids comprise the sequences set forth in SEQ ID NOS:4, 7, 10, and 12, respectively, or portions thereof, or nucleotide sequences encoding, in whole or in part, a CD24*, Set*, GluT1*, or 115392* protein (*e.g.*, a protein comprising the amino acid sequence of SEQ ID NOS:5, 8, 11 or 13, respectively, or portions thereof). The present invention also discloses purified nucleic acids comprised of at least six (6) nucleotides (*i.e.*, a hybridizable portion) of a *CD24**, *Set**, *GluT1**, or *115392** nucleic acid sequence. In a preferred embodiment of the present invention, the nucleic acids consisting of at least six (6) nucleotides of a *Set** or *GluT1** nucleotide sequence are comprised of those sequences which span the splice junctions specific to Set* and GluT1*, respectively (*i.e.*, the portion of the *Set** or *GlutT1** nucleotide sequence containing the sequence at which the 5'- and 3'-splice sites used in processing *Set** and *GluT1** mRNA but not the *Set* mRNA or *GluT1* mRNA were joined). By way of example, but not of limitation, a nucleic acid sequence consisting of nucleotides both 5'- and 3'-of the splice site at nucleotide numbers 212-213 of the *Set** nucleotide sequence as depicted in Figure 4 [SEQ ID NO:7] or nucleotide numbers 860-861 of the *GluT1** nucleotide sequence as depicted in Figure 6 [SEQ ID NO:10]. In other embodiments, the nucleic acids consist of at least 25 (continuous) nucleotides, 50 nucleotides, 100 nucleotides, 150 nucleotides, or 200 nucleotides *of* a *CD24**, *Set**, *GluT1**, or *115392** sequence, or a full-length *CD24**, *Set**, *GluT1**, or *115392** coding sequence. In yet another embodiment, the nucleic acid sequences are smaller than 35, 200 or 500 nucleotides in length. The nucleic acid sequences may be either single or double-stranded.

The invention also relates to nucleic acids hybridizable to or complementary to the foregoing sequences, in particular the invention provides the inverse complement to nucleic acids hybridizable to the foregoing sequences (*i.e.*, the inverse complement of a nucleic acid strand has the complementary sequence running in reverse orientation to the strand so that the inverse complement would hybridize without mismatches to the nucleic acid strand; thus, for example, where the coding strand is hybridizable to a nucleic acid with no mismatches between the coding strand and the hybridizable strand, then the inverse complement of the hybridizable strand is identical to the coding strand). In specific embodiments of the present invention, nucleic acids are provided which comprise a sequence complementary to (specifically are the inverse

complement of) at least 10, 25, 50, 100, or 200 nucleotides or the entire coding region of a *CD24*, Set*, GluT1**, or *115392** gene.

In a specific embodiment of the present invention, a nucleic acid which is hybridizable to a *CD24*, Set*, GluT1**, or *115392** nucleic acid (*e.g.*, having sequence SEQ ID NOS:4, 7, 10, and 12 respectively), or to a nucleic acid encoding a CD24*, Set*, GluT1*, or 115392* derivative, (or the complement of the foregoing) under conditions of low stringency is provided. By way of example, and not of limitation, procedures using such conditions of low stringency are as follows. See *e.g.*, Shilo and Weinberg, 1981. *Proc. Natl. Acad. Sci. USA* 78:6789-6792. Filters containing DNA are pre-hybridized for 6 hours at 40°C in a solution containing 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.1% PVP, 0.1% Ficoll, 1% BSA, and 500 μg/ml denatured salmon sperm DNA. Hybridizations are carried out in the same solution with the following modifications: 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 μg/ml salmon sperm DNA, 10% (wt/vol) dextran sulfate, and 5-20 x 10^6 cpm ^{32}P-labeled probe. Filters are incubated in hybridization mixture for 18-20 hours at 40°C, and then washed for 1.5 hours at 55°C in a solution containing 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS. The wash solution is replaced with fresh solution and incubated an additional 1.5 hours at 60°C. Filters are blotted dry and exposed to X-ray film for autoradiography. If necessary, filters are washed for a third time at 65-68°C and re-exposed to X-ray film. Other conditions of low stringency which may be used are well known in the art.

In another specific embodiment, a nucleic acid which is hybridizable to a *CD24*, Set*, GluT1**, or *115392** nucleic acid or a nucleic acid encoding a CD24*, Set*, GluT1* or 115392* derivative (or a complement of the foregoing) under conditions of moderate stringency is provided. For example, but not limited to, procedures using such conditions of moderate stringency are as follows: Filters containing DNA are pretreated for 6 hours at 55°C in a solution containing 6X SSC, 5X Denhardt's solution, 0.5% SDS and 100 μg/ml denatured salmon sperm DNA. Hybridizations are carried out in the same solution with 5-20 x 10^6 cpm ^{32}P-labeled probe. Filters are incubated in hybridization mixture for 18-20 hours at 55°C, and then washed twice for 30 minutes at 60°C in a solution containing 1X SSC and 0.1% SDS. Filters are blotted dry and exposed to X-ray film for autoradiography. Washing of filters is done at 37°C for 1 hour in a solution containing 2X SSC, 0.1% SDS. Other conditions of moderate stringency which may be used are well-known in the art.

In another specific embodiment, a nucleic acid which is hybridizable to a *CD24*, Set*, GluT1** or *115392** nucleic acid (or a complement of the foregoing) or to a nucleic acid encoding a CD24*, Set* GluT1* or 115392* derivative under conditions of high stringency is provided.

By way of example and not limitation, procedures using such conditions of high stringency are as follows: pre-hybridization of filters containing DNA is carried out for 8 hours to overnight at 65°C in buffer composed of 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, and 500 µg/ml denatured salmon sperm DNA. Filters are hybridized for 48 hours at 65°C in pre-hybridization mixture containing 100 µg/ml denatured salmon sperm DNA and 5-20 x 10^6 cpm of ^{32}P-labeled probe. Washing of filters is done at 37°C for 1 hour in a solution containing 2X SSC, 0.01% PVP, 0.01% Ficoll, and 0.01% BSA. This is followed by a wash in 0.1X SSC at 50°C for 45 minutes prior to autoradiography. Other conditions of high stringency which may be used are well known in the art.

Nucleic acids encoding derivatives and analogs of CD24*, Set*, GluT1*, or 115392* proteins (see this Section, *supra*), and *CD24*, Set*, GluT1**, or *115392** antisense nucleic acids are additionally disclosed. As is readily apparent, as utilized herein, a "nucleic acid encoding a fragment or portion of CD24*, Set*, GluT1* or 115392*" shall be construed as referring to a nucleic acid encoding only the recited fragment or portion of CD24*, Set*, GluT1*, or 115392* protein, and not the other contiguous portions of the CD24*, Set*, GluT1*, or 115392* as a continuous sequence.

Sequences within the 3'-non-translated regions of the *CD24*, *Set*, and *GluT1* mRNAs were identified as encoding a protein or proteins that interact with RB using an improved, modified version of the yeast two hybrid system. The present inventors have identified interacting proteins encoded by nucleotide sequences identical to the portion of the nucleotide sequence of *CD24* that is 3'- to base 689, of the nucleotide sequence of *Set* that is 3'- to base 2076, and of the nucleotide sequence of *GluT1* that is 3' to base 3408.

Within nucleotide sequences, potential open reading frames can be identified using the N.C.B.I. BLAST program "ORF Finder" available to the public. Because all known protein translation products are at least 60 amino acids or longer (see *e.g.*, Creighton, 1992. *Proteins, 2nd Ed.*, W.H. Freeman and Co., New York), only those ORFs potentially encoding a protein of 60 amino acids or more are considered. Furthermore, if an initiation methionine codon (ATG) and a translational stop codon (TGA, TAA, or TGA) are identified, then the boundaries of the protein are defined. Other potential proteins include any open reading frames that extend to the 5'-terminus of the nucleotide sequence, in which case the open reading frame predicts the C-terminal portion of a longer protein. Similarly, any open reading frame that extends to the 3'-terminus of the nucleotide sequence predicts the N-terminal portion of a longer protein. This methodology was used to identify open reading frames encoding the interactants CD24* (Figure

1; SEQ ID NO: 1) and 115392* (Figure 7; SEQ. ID. NO:12) from the CD24 and the 115392 nucleotide sequences.

That the nucleotide sequence encoding the portions of the CD24*, Set* and GluT1* identified as interacting with RB are 3' to the translational stop codon of the nucleotide sequence encoding CD24, Set and GluT1, respectively, indicates that CD24*, Set* and GluT1* could be encoded by mRNAs resulting from splicing of unprocessed *CD24*, *Set* and *GluT1* gene transcripts at splice sites other than the splice sites used in processing the known *CD24*, *Set* and *GluT1* mRNAs. The *Set** and *GluT1** sequences were determined by identifying alternate 5' and 3'-splice sites in the *Set* and *GluT1* sequence. Alternate splice sites within the *CD24* sequence were not detected.

Determination of 5' and 3' splice points for protein splice variants can be performed by any method known within the art. By way of example, but not of limitation, the 5' and 3' splice points may be ascertained by the following methodology:

(1) Potential 5'-splice sites can be identified in the coding sequence of the known protein (*e.g.*, Set and GluT1). The sequence of 5'-splice sites has an invariant GT sequence at the start of the intron, and the remaining bases are not invariant, but the preferred consensus sequence is AG:GTAAGT, with the colon indicating the splice point (see Padgett, *et al.*, 1984. *Ann. Rev. Biochem.* 55:1119-1150). Potential splice sites can be identified in order of the number of residues matching this consensus sequence, requiring at a minimum, the invariant GT and 4/6 matches in the other consensus bases. In cases where no potential 5'-splice sites are identified, the required 4/6 matches in the consensus bases outside of the GT sequence is relaxed to 3/6 matches.

(2) Potential 3'- intron:exon splice sites can also be identified based on the consensus analysis described by Padgett, *et al.* (1984. *Ann. Rev. Biochem.* 55:1119-1150). The 3'-intron:exon splice site must have an AG sequence immediately 5'- to the splice site (denoted as "AG:") and the base just 5'- (preceding) the AG: sequence must be a C or a T. Then, the nucleotides which are 5 to 14 nucleotides 5'- of the last intronic G base can contain at most two non-T, non-C bases. To identify such a potential 3' intron:exon splice site, the sequence between potential 5'-splice sites and the start of the nucleotide sequence encoding the detected interacting protein region can be scanned for the invariant AG: sequence, where the base preceding the invariant region must be a C or T. Potential sites are then examined to determine if they meet the requirement of at least 8 of 10 C or T bases in positions -5 to -14 from the splice site.

(3) Based upon the translational frame of the known protein and each predicted 5'-splice site, compatible translational frames for successful splicing can be defined for potential 3'-splice sites. Nucleotide sequences can be analyzed by a number of nucleotide sequence analysis programs available in the art to define possible protein translation products (*e.g.*, ORF Finder in the N.C.B.I. BLAST Package (N.C.B.I., Bethesda, MD)). Translation in the three forward translation frames defines possible open reading frames (contiguous spans of codons for amino acids without the presence of a stop codon). Only those 3'- sites that match the necessary translational frame of a 5'-splice junction are retained. Unmatched 5'- or 3'-splice sites are

eliminated. In cases where no ideal 3'-splice site match is found, sites containing three non-C, non-T bases upstream of the splice site are then examined.

(4) For each possible 5':3' splice site pair, a search for a mammalian branch point consensus sequence is performed (see Reed and Maniatis, 1988. *Genes Dev.* 2:1268-1276). The branch point is identified by the consensus sequence T/CNCTGAC to which 5 of the 6 defined bases must match and the consensus sequence must be 20-60 nucleotides 5'- to the 3'-splice site. Though not absolutely required for pre mRNA splicing, the efficiency of splicing is related to the presence of the consensus sequence. Thus, 5':3' splice site pairs with a branch point consensus sequence are retained over splice site pairs that do not have a branch point consensus sequence. Each 5':3' splice site pair that remains defines a new splice variant of the known protein. Herein, a splice variant is named after the known protein, followed by a *, and if more than one splice variant for a given gene is identified, the name is followed by a * and number.

(5) New splice variant proteins must encode at least 60 amino acid residues to constitute a viable *in vivo* product. Further, the 3'-terminus of slice variants must, by definition, extend into the identified interacting sequence.

The amino acid and nucleotide sequences for the splice variants of Set and GluT1, named Set* and GluT1*, respectively, and depicted in Figures 4 and 6, respectively, were determined *in silico* as described above and as exemplified in the Specific Examples Section, *infra*. For *Set**, a 5'-splice site was identified at nucleotides 214-221 of the *Set* nucleotide sequence (indicated by "**B**" on the *Set* nucleotide sequence in Figure 3, with the last base of the first exon being nucleotide number 215, as indicated by the arrow in callout "**B**" in Figure 3, and a 3'-splice site was identified at nucleotides 2061 to 2074 of the *Set* nucleotide sequence (indicated by callout "**E**" on the *Set* nucleotide sequence in Figure 3, with the first base of the second exon being nucleotide number 2074, as indicated by the arrow in Figure 3, callout at "**E**"). The translation stop codon of *Set** was identified as nucleotides 2211 to 2213 of the *Set* nucleotide sequence indicated as callout "E" in Figure 3. The branch point consensus region for *Set** splicing was identified at nucleotides 2053 to 2060 of the *Set* nucleotide sequence as depicted in Figure 3 (callout "**D**"). The final splice variant nucleotide sequence of Set* is depicted in Figure 4 [SEQ ID NO:8].

For *GluT1**, a 5'-splice site was identified at nucleotides 1037-1044 of the *GluT1* sequence, as depicted in Figure 5 [SEQ ID NO:9] and as indicated by callout "**B**", with the last base of the first exon being nucleotide number 1038, as indicated by the arrow in Figure 5 (callout "**B**"), and a 3'-splice site was identified at nucleotides 3366-3380 of the *GluT1* nucleotide sequence, as depicted in Figure 5 (callout "**E**"), with the first base of the second exon being nucleotide number 3380, as indicated by the arrow in Figure 5 (callout "**E**"). The branch point site associated with this 3'-splice site is at nucleotides 3301-3307 of the *GluT1* sequence in

Figure 5 and is indicated as callout "**D**", and the translation stop codon for *GluT1** is nucleotides 3488-3470 of the *GluT1* nucleotide sequence, indicated as callout "**H**" in Figure 5.

Any method available in the art can be used to obtain a full length (*i.e.*, encompassing the entire coding region) cDNA clone encoding CD24*, Set*, GluT1*, or 115392*. In particular, the polymerase chain reaction (PCR) can be used to amplify sequences defined *in silico* from a cDNA library. Oligonucleotide primers that hybridize to sequences at the 3'- and 5'-termini of the identified sequences can be used as primers to amplify by PCR sequences from a nucleic acid sample (cDNA or DNA), preferably a cDNA library, from an appropriate source (*e.g.*, the sample from which the initial cDNA library for the improved, modified yeast two hybrid assay fusion population was derived).

PCR can be carried out, for example, by use of a Perkin-Elmer Cetus thermal cycler and *Taq* polymerase (Gene Amp™). The DNA being amplified can include genomic DNA or cDNA sequences from any eukaryotic species. One may choose to synthesize several different degenerate primers, for use in the PCR reactions. It is also possible to vary the stringency of hybridization conditions used in priming the PCR reactions, to amplify nucleic acid homologs (*e.g.*, to obtain CD24*, Set*, GluT1*, or 115392* sequences from species other than humans, or to obtain human sequences with homology to CD24*, Set*, GluT1*, or 115392*) by allowing for greater or lesser degrees of nucleotide sequence similarity between the known nucleotide sequence and the nucleic acid homolog being isolated. For cross species hybridization, low stringency conditions are preferred. For same species hybridization, moderately stringent conditions are preferred.

After successful amplification of the nucleic acid containing all or a portion of the CD24*, Set*, GluT1*, or 115392* sequence, that segment may be molecularly cloned and sequenced, and utilized as a probe to isolate a complete cDNA or genomic clone. This, in turn, will permit the determination of the gene's complete nucleotide sequence, the analysis of its expression, and the production of its protein product for functional analysis, as described *infra*. In this fashion, the nucleotide sequences of the entire *CD24*, Set*, GluT1*, or 115392** genes, as well as additional genes encoding CD24*, Set*, GluT1*, or 115392* proteins and analogs may be identified.

Any eukaryotic cell potentially can serve as the nucleic acid source for the molecular cloning of the *CD24*, Set*, GluT1*, or 115392** gene. The nucleic acids can be isolated from vertebrates, including: mammalian, human, porcine, bovine, feline, avian, equine, canine, as well as additional primate sources, insects, plants, and the like. The DNA may be obtained by standard procedures known in the art from cloned DNA (*e.g.*, a DNA "library"), by chemical

synthesis, by cDNA cloning, or by the cloning of genomic DNA, or fragments thereof, purified from the desired cell (see *e.g.*, Sambrook, *et al.*, 1989. *Molecular Cloning, A Laboratory Manual, 2d Ed.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York; Glover, D.M. (ed.), 1985. *DNA Cloning: A Practical Approach*, MRL Press, Ltd., Oxford, U.K. Vol. I, II). The clones derived from such genomic DNA may contain regulatory and intronic DNA regions in addition to coding regions; clones derived from cDNA will contain only exon sequences. Whatever the source, the gene should be molecularly cloned into a suitable vector for propagation of the gene.

In the molecular cloning of the gene from genomic DNA, DNA fragments are generated, some of which will encode the desired gene. The DNA may be cleaved at specific sites using various restriction enzymes. Alternatively, one may use DNase in the presence of manganese to fragment the DNA, or the DNA can be physically sheared (*e.g.*, by sonication). The linear DNA fragments can then be separated according to size by standard techniques including, but not limited to, agarose and/or polyacrylamide gel electrophoresis, and column chromatography.

Once the DNA fragments are generated, identification of the specific DNA fragment containing the desired gene may be accomplished in a number of ways. For example, a portion of the *CD24*, Set*, GluT1*, or 115392** (of any species) gene (*e.g.*, a PCR amplification product obtained as described above, or an oligonucleotide having a sequence of a portion of the known nucleotide sequence) or its specific RNA, or a fragment thereof, may be purified and labeled, and the generated DNA fragments may be screened by nucleic acid hybridization to the labeled probe (see *e.g.*, Benton and Davis, 1977. *Science* 196:180-182; Grunstein and Hogness, 1975. *Proc. Natl. Acad. Sci. U.S.A.* 72:3961-3964). Those DNA fragments with substantial homology to the probe will hybridize. It is also possible to identify the appropriate fragment by restriction enzyme digestion(s) and comparison of fragment sizes with those expected according to a known restriction map if such is available, or by DNA sequence analysis and comparison to the known nucleotide sequence of *CD24*, Set*, GluT1*, or 115392**. Further selection can be carried out on the basis of the properties of the gene. Alternatively, the presence of the gene may be detected by assays based on the physical, chemical, or immunological properties of its expressed product. For example, cDNA clones, or DNA clones which hybrid-select the proper mRNAs, can be selected which produce a protein that, for example, possesses similar or identical electrophoretic migration, isoelectric focusing behavior, proteolytic digestion maps, or antigenic properties or ability to bind RB, as is known for CD24*, Set*, GluT1*, or 115392*. In addition, if an anti-CD24*, anti-Set*, anti-GluT1*, or anti-115392* antibody is available, the protein may

be identified by binding of labeled antibody to the putatively CD24*, Set*, GluT1*, or 115392* synthesizing clones, in an enzyme-linked immunosorbent assay (ELISA)-type procedure.

An alternative to isolating the *CD24*, Set*, GluT1*, or 115392** cDNA includes, but is not limited to, chemically synthesizing the gene sequence itself from a known sequence. Other methods are possible and within the scope of the present invention.

The identified and isolated nucleic acids can then be inserted into an appropriate cloning vector. A large number of vector-host systems known in the art may be used. Possible vectors include, but are not limited to, plasmids or modified viruses, but the vector system must be compatible with the host cell used. Such vectors include, but are not limited to, bacteriophages such as lambda derivatives, or plasmids such as pBR322 or pUC plasmid derivatives or the pBlueScript vector (Stratagene, La Jolla, CA). Insertion into a cloning vector can, for example, be accomplished by ligating the DNA fragment into a cloning vector which has complementary cohesive termini. However, if the complementary restriction sites used to fragment the DNA are not present in the cloning vector, the ends of the DNA molecules may be enzymatically modified to ensure compatibility. Alternatively, any site desired may be produced by ligating nucleotide sequences (*e.g.*, linkers) onto the DNA termini; these ligated linkers may comprise specific chemically synthesized oligonucleotides encoding restriction endonuclease recognition sequences. In an alternative method, the cleaved vector and the *CD24*, Set*, GluT1*, or 115392** gene may be modified by homopolymeric tailing. Recombinant molecules can be introduced into host cells via transformation, transfection, infection, electroporation, etc., so that many copies of the gene sequence are generated.

In an alternative method, the desired gene may be identified and isolated after insertion into a suitable cloning vector in a "shot gun" approach. Enrichment for the desired gene, for example, by size fractionation, can be done before insertion into the cloning vector.

In specific embodiments, transformation of host cells with recombinant DNA molecules that incorporate the isolated *CD24*, Set*, GluT1*, or 115392** gene, cDNA, or synthesized DNA sequence enables generation of multiple copies of the gene. Thus, the gene may be obtained in large quantities by growing transformants, isolating the recombinant DNA molecules from the transformants and, when necessary, retrieving the inserted gene from the isolated recombinant DNA.

The *CD24*, Set*, GluT1*, or 115392** sequences provided by the instant invention include those nucleotide sequences encoding substantially the same amino acid sequences as found in native CD24*, Set*, GluT1*, or 115392* proteins, and those encoded amino acid sequences with functionally equivalent amino acids, as well as those encoding other CD24*,

Set*, GluT1*, or 115392* derivatives or analogs, as described in Section (I), *supra,* for CD24*, Set*, GluT1*, or 115392* derivatives and analogs

(III) Antibodies to RB•RB-IP Complexes and CD24*, Set*, GluT1* and 115392* Proteins

As disclosed by the present invention, the RB•RB-IP complexes (*e.g.*, RB complexed with RN-tre, CAS, IP-30, L6, IK, cyclophilin A, Zap 3, Nlk1(Nek 2), 17 beta-HSD6, CD24*, Set*, GluT1*, or 115392*), or fragments, derivatives or homologs thereof, or CD24*, Set*, GluT1*, or 115392* proteins, and fragments, homologs and derivatives thereof, may be utilized as immunogens to generate antibodies which immunospecifically bind such immunogens. Such antibodies include, but are not limited to polyclonal, monoclonal, chimeric, and single chain antibodies, F_{ab} fragments, and F_{ab} expression libraries. In a specific embodiment, antibodies to complexes of human RB and a human RB-IP are produced. In another embodiment, complexes formed from fragments of an RB and an RB-IP, where the fragments contain the protein domain that interacts with the other member of the complex, are used as immunogens for antibody production. In another specific embodiment, CD24*, Set*, GluT1*, or 115392* proteins or fragments, derivatives, or homologs thereof are used as immunogens.

Various procedures known in the art may be used for the production of polyclonal antibodies to a RB•RB-IP complex, or derivative or analog thereof, or CD24*, Set*, GluT1*, or 115392* protein, or derivative, fragment or analog thereof.

For production of the antibody, various host animals can be immunized by injection with the native RB•RB-IP complex, or CD24*, Set*, GluT1*, or 115392* protein, or a synthetic version, or a derivative of the foregoing, such as a cross-linked RB•RB-IP. Such host animals include, but are not limited to, rabbits, mice, rats, etc. Various adjuvants can be used to increase the immunological response, depending on the host species, and include, but are not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, and potentially useful human adjuvants such as bacille Calmette-Guerin (BCG) and *Corynebacterium parvum*.

For preparation of monoclonal antibodies directed towards an RB•RB-IP complex or CD24*, Set*, GluT1*, or 115392*, or derivatives, fragments or analogs thereof, any technique that provides for the production of antibody molecules by continuous cell lines in culture may be used. Culture techniques include, but are not limited to: (*i*) the hybridoma technique (see Kohler & Milstein, 1975. *Nature* 256:495-497); (*ii*) the trioma technique see Rosen, *et al.*, 1977. *Cell* 11:139-147); (*iii*) the human B-cell hybridoma technique (see Kozbor, *et al.*, 1983. *Immunology*

Today 4:72) and the EBV hybridoma technique to produce human monoclonal antibodies (see *e.g.*, Cole, *et al.*, 1985. In: *Monoclonal Antibodies and Cancer Therapy* (Alan R. Liss, Inc.). In an additional embodiment of the invention, monoclonal antibodies can be produced in germ-free animals utilizing a recently developed technology (see *e.g.*, PCT Patent Publication U590/02545).

In a specific embodiment of the present invention, human antibodies may be used and can be obtained by using human hybridomas (see *e.g.*, Cole, *et al.*, 1983. *Proc. Natl. Acad. Sci. USA* 80:2026-2030) or by transforming human B cells with EBV virus *in vitro* (see *e.g.*, Cole, *et al.*, 1985. In: *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96). In fact, according to the invention, techniques developed for the production of "chimeric antibodies" (see *e.g.*, Morrison, *et al.*, 1984. *Proc. Natl. Acad. Sci. USA* 81: 6851-6855; Takeda, *et al.*, 1985. *Nature* 314:452-454) by splicing the genes from a mouse antibody molecule specific for the RB•RB-IP complex or CD24*, Set*, GluT1*, or 115392* protein, together with genes from a human antibody molecule of appropriate biological activity, may be used. The production of such antibodies are within the scope of the present invention.

In another specific embodiment of the present invention, techniques described for the production of single-chain antibodies (see *e.g.*, U.S. Patent 4,946,778) may be adapted to produce RB•RB-IP complex-specific and CD24*-, Set*-, GluT1*-, and 115392*-specific single-chain antibodies. An additional embodiment of the invention utilizes the techniques described for the construction of F_{ab} expression libraries (see *e.g.*, Huse, *et al.*, 1989. *Science* 246:1275-1281) to allow rapid and easy identification of monoclonal F_{ab} fragments with the desired specificity for the RB•RB-IP complexes, or individual CD24*, Set*, GluT1*, or 115392* proteins, derivatives or analogs thereof. Non-human antibodies can be "humanized" by known methods (see *e.g.*, U.S. Patent No. 5,225,539).

Antibody fragments that contain the idiotypes of RB•RB-IP complexes or of CD24*, Set*, GluT1*, or 115392* proteins can be generated by techniques known in the art. For example, such fragments include, but are not limited to: the $F_{(ab')2}$ fragment which can be produced by pepsin digestion of the antibody molecule; the $F_{ab'}$ fragments that can be generated by reducing the disulfide bridges of the $F_{(ab')2}$ fragment; the F_{ab} fragments that can be generated by treating the antibody molecular with papain and a reducing agent; and F_v fragments.

In the production of antibodies, screening for the desired antibody can be accomplished by various techniques known within the art (enzyme-linked immunosorbent assay (ELISA)). To select antibodies specific to a particular domain of the RB•RB-IP complex, or CD24*, Set*, GluT1*, or 115392* protein, one may assay generated hybridomas for a product that binds to the

fragment of the RB•RB-IP complex, or CD24*, Set*, GluT1*, or 115392* protein, that contains such a domain. For selection of an antibody that specifically binds an RB•RB-IP complex but which does not specifically bind to the individual proteins of the RB•RB-IP complex, one can select on the basis of positive binding to the complex *in toto* and a lack of binding to the individual RB and RB-IP proteins. Additionally, in specific embodiments, antibodies are generated that immunospecifically recognize the portion of Set* or GluT1* having an amino acid sequence that differs from the Set or GluT1 amino acid sequence and that do not immunospecifically recognize Set or GluT1. In addition, antibodies specific to a domain of the RB•RB-IP complex are also disclosed, as are antibodies to specific domains of CD24*, Set*, GluT1*, and 115392* proteins.

The aforementioned antibodies may also be used in methods known within the art relating to the localization and/or quantitation of RB•RB-IP complexes and CD24*, Set*, GluT1*, and 115392* proteins of the invention (*e.g.*, for imaging these proteins, measuring levels thereof in appropriate physiological samples, in diagnostic methods, and the like).

In yet another embodiment of the present invention, anti-RB•RB-IP complex antibodies and fragments thereof, or anti-CD24*, anti-Set*, anti-GluT1*, and anti-115392*, or fragments thereof, containing the binding domain, are utilized as Therapeutic moieties (see *infra*).

(IV) The Utilization of RB•RB-IP Complexes and Nucleic Acids and CD24*, Set*, GluT1* and 115392* Proteins and Nucleic Acids in Diagnostic, Prognostic and Screening Methodologies

RB•RB-IP complexes (particularly RB complexed with RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, and 115392*), or CD24*, Set*, GluT1*, and 115392* proteins, may be markers of normal: (*i*) physiological processes (*e.g.*, cell cycle control, cellular differentiation and apoptosis); (*ii*) response to viral infection; (*iii*) intracellular signal transduction; (*iv*) transcriptional control; and (*v*) pathophysiological processes (*e.g.*, hyperproliferative disorders including tumorigenesis and tumor spread, degenerative disorders including neurodegenerative diseases, virus infection), and thus possess diagnostic utility. Further, definition of particular groups of patients with elevations or deficiencies of an RB•RB-IP complex, or CD24*, Set*, GluT1*, or 115392* proteins, can lead to new nosological classifications of diseases, furthering diagnostic ability.

Detecting levels of RB•RB-IP complexes, or individual proteins that have been shown to form complexes with RB, or the CD24*, Set*, GluT1*, or 115392* proteins, or detecting levels of mRNAs encoding components of the RB•RB-IP complexes, or CD24*, Set*, GluT1*, and

115392* proteins, may be used in prognosis, to follow the course of disease states, to follow therapeutic response, and the like.

RB•RB-IP complexes and the individual components of the RB•RB-IP complexes (*e.g.*, RB, RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, and 115392*), and derivatives, analogs and subsequences thereof, RB and/or RB-IP, and CD24*, Set*, GluT1*, and 115392* nucleic acids (and sequences complementary thereto), and anti-RB•RB-IP complex antibodies and antibodies directed against the individual components that can form RB•RB-IP complexes, and anti-CD24*, anti-Set*, anti-GluT1*, and anti-115392* antibodies, have uses in diagnostics. Such molecules can be used in assays, such as immunoassays, to detect, prognose, diagnose, or monitor various conditions, diseases, and disorders, and treatment thereof, characterized by aberrant levels of RB•RB-IP complexes, or CD24*, Set*, GluT1*, or 115392* proteins.

More specifically, by way of example and not of limitation, such an immunoassay is performed by use of a method comprising contacting a sample derived from a patient with an anti-RB•RB-IP complex antibody, or anti-CD24*, anti-Set*, anti-GluT1*, or anti-115392* antibody under conditions such that immunospecific binding can occur, and detecting or measuring the amount of any immunospecific binding by the antibody. In a specific aspect, such binding of antibody, in tissue sections, can be used to detect aberrant RB•RB-IP complex formation, or aberrant CD24*, Set*, GluT1*, or 115392* protein localization, or aberrant (*e.g.*, high, low or absent) levels of RB•RB-IP complex or complexes, or CD24*, Set*, GluT1*, and 115392* proteins. In a specific embodiment of the present invention, an antibody against an RB•RB-IP complex can be used to assay a patient tissue or serum sample for the presence of the RB•RB-IP complex, where an aberrant level of said complex is an indication of a diseased condition. In another embodiment, an antibody against CD24*, Set*, GluT1*, or 115392* can be used to assay a patient tissue or serum sample for the presence of CD24*, Set*, GluT1*, or 115392* where an aberrant level of CD24*, Set*, GluT1*, and 115392* is an indication of a diseased condition. The term "aberrant levels," as utilized herein, is defined as meaning increased or decreased levels relative to that present, or a standard level representing that present, in an analogous sample from a portion of the body, or from a subject not having the specific disorder.

The immunoassays which may be utilized include, but are not limited to: competitive and non-competitive assay systems using techniques such as Western blots, radioimmunoassays (RIA), enzyme-linked immunosorbent assay (ELISA), "sandwich" immunoassays, immunoprecipitation assays, precipitin reactions, gel diffusion precipitin reactions,

immunodiffusion assays, agglutination assays, complement-fixation assays, immunoradiometric assays, fluorescent immunoassays, and protein A immunoassays, and other similar methodologies.

Nucleic acids encoding the various protein constituents of the RB•RB-IP complexes and encoding the CD24*, Set*, GluT1*, and 115392* proteins, and related nucleotide sequences and subsequences, including complementary sequences, can also be used in hybridization assays. The RB and/or RB-IP nucleotide sequences, or subsequences thereof, comprising about at least six (6) nucleotides, may be used as hybridization probes. Hybridization assays may then be utilize to detect, prognose, diagnose, or monitor conditions, disorders, or disease states associated with aberrant levels of the mRNAs encoding the constituent proteins of an RB•RB-IP complex, or CD24*, Set*, GluT1*, and 115392* protein, as described *supra*. Specifically, such a hybridization assay is carried out by a method comprising contacting a sample containing nucleic acid with a nucleic acid probe capable of hybridizing to RB or an RB-IP DNA or RNA, under conditions such that hybridization can occur, and detecting or measuring any resulting hybridization. In a preferred aspect of the present invention, the hybridization assay is carried out using nucleic acid probes capable of hybridizing to RB and to a binding partner (*i.e.*, IP) of to measure concurrently the expression of both members of an RB•RB-IP complex. In another preferred embodiment, the expression of mRNAs encoding CD24*, Set*, GluT1*, or 115392* is quantitatively ascertained.

In specific embodiments of the present invention, diseases and disorders involving or characterized by aberrant levels of RB•RB-IP complexes may be diagnosed, or their suspected presence can be screened for, or a predisposition to develop such disorders can be detected, by detecting aberrant levels of RB•RB-IP complexes, or non-complexed RB and/or an RB-IP proteins or nucleic acids or functional activity including, but not restricted to: (*i*) binding to an interacting partner or (*ii*) detecting mutations in RB and/or an RB-IP RNA, DNA or protein (*e.g.*, translocations, truncations, changes in nucleotide or amino acid sequence relative to wild-type RB and/or RB-IP) that cause increased or decreased expression or activity of a RB•RB-IP complex and/or RB and/or protein that binds to RB. Such diseases and disorders include, but are not limited to, those described in Section (V).

By way of example and not of limitation, levels of RB•RB-IP complexes and the individual components of RB•RB-IP complexes may be detected by immunoassay; levels of *RB* and/or *RB-IP* RNA may be detected by hybridization assays (*e.g.*, Northern blots, dot blots); binding of RB to an RB-IP may be measured by binding assays commonly known in the art, translocations and point mutations in *RB* and/or in genes encoding RB-IPs may be detected by

Southern blotting, RFLP analysis, PCR using primers that preferably generate a fragment spanning at least most of the *RB* and/or *RB-IP* gene, sequencing of the *RB* and/or *RB-IP* genomic DNA or cDNA obtained from the patient, and the like.

Assays which are well-known within the art (*e.g.*, immunoassays, nucleic acid hybridization assays, activity assays, etc.) may be used to determine whether one or more particular RB•RB-IP complexes are present at either increased or decreased levels, or are absent, in samples from patients suffering from a particular disease or disorder, or having a predisposition to develop such a disease or disorder as compared to the levels in samples from subjects not having such a disease or disorder.

Additionally, these aforementioned assays may also be used to determine whether the ratio of the RB•RB-IP complex, or CD24*, Set*, GluT1*, or 115392*, to the non-complexed components of the RB•RB-IP complex, or CD24*, Set*, GluT1*, or 115392*, (*i.e.*, RB and/or the specific RB-IP in the complex of interest, is increased or decreased in samples from patients suffering from a particular disease or disorder, or having a predisposition to develop such a disease or disorder, as compared to the ratio in samples from subjects not having such a disease or disorder).

In the event that levels of one or more particular RB•RB-IP complexes, or CD24*, Set*, GluT1*, or 115392*, are determined to be increased in patients suffering from a particular disease or disorder, or having a predisposition to develop such a disease or disorder, then the particular disease or disorder, or predisposition for a disease or disorder, can be diagnosed, have prognosis defined for, be screened for, or be monitored by detecting increased levels of the one or more RB•RB-IP complexes, , or CD24*, Set*, GluT1*, or 115392*, the mRNA that encodes the members of the one or more particular RB•RB-IP complexes, or CD24*, Set*, GluT1*, or 115392*, or RB•RB-IP complex, or CD24*, Set*, GluT1*, or 115392*, functional activity.

Accordingly, in a specific embodiment of the present invention, diseases and disorders involving increased levels of one or more RB•RB-IP complexes, or CD24*, Set*, GluT1*, or 115392*, may be diagnosed, or their suspected presence can be screened for, or a predisposition to develop such disorders can be detected, by detecting increased levels of the one or more RB•RB-IP complexes, , or CD24*, Set*, GluT1*, or 115392*, the mRNA encoding both members of the complex, or complex functional activity, or by detecting mutations in RB or the RB-IP (*e.g.*, translocations in nucleic acids, truncations in the gene or protein, changes in nucleotide or amino acid sequence relative to the native, wild-type RB or RB-IP) which function to stabilize or increase RB•RB-IP complex, or CD24*, Set*, GluT1*, or 115392*, formation.

In the event that levels of one or more particular RB•RB-IP complexes, or CD24*, Set*, GluT1*, or 115392*, are determined to be decreased in patients suffering from a particular disease or disorder or having a predisposition to develop such a disease or disorder, then the particular disease or disorder or predisposition for a disease or disorder can be diagnosed, have its prognosis determined, be screened for, or be monitored by detecting decreased levels of the one or more RB•RB-IP complexes, or CD24*, Set*, GluT1*, or 115392*, the mRNA that encodes the members of the particular one or more RB•RB-IP complexes, or CD24*, Set*, GluT1*, or 115392*, or RB•RB-IP complex, or CD24*, Set*, GluT1*, or 115392*, functional activity.

Accordingly, in a specific embodiment of the present invention, diseases and disorders involving decreased levels of one or more RB•RB-IP complexes, or CD24*, Set*, GluT1*, or 115392*, can be diagnosed, or their suspected presence can be screened for, or a predisposition to develop such disorders can be detected, by detecting decreased levels of the one or more RB•RB-IP complexes, , or CD24*, Set*, GluT1*, or 115392*,the mRNA encoding the members of the one or more complexes, or complex functional activity, or by detecting mutations in RB or the RB-IP (*e.g.*, translocations in nucleic acids, truncations in the gene or protein, changes in nucleotide or amino acid sequence relative to wild-type RB or the RB-IP) which function to inhibit or reduce RB•RB-IP complex formation.

Also disclosed by the present invention are methodologies for the detection of an RB•RB-IP complex, or CD24*, Set*, GluT1*, or 115392* protein, in cell culture models that express particular RB•RB-IP complexes or CD24*, Set*, GluT1*, or 115392* proteins, or derivatives thereof, for the purpose of characterizing or preparing RB•RB-IP complexes, or CD24*, Set*, GluT1*, or 115392* proteins for harvest. This aforementioned embodiment includes, but is not limited to, cell sorting of prokaryotes (see *e.g.*, Davey and Kell, 1996. *Microbiol. Rev.* 60: 641-696); primary cultures and tissue specimens from eukaryotes, including mammalian species such as human (see *e.g.*, Steele, *et al.*, 1996. *Clin. Obstet. Gynecol.* 39:801-813) and continuous cell cultures (see *e.g.*, Orfao and Ruiz-Arguelles, 1996. *Clin. Biochem.* 29:5-9). In addition, such isolation methodologies may also be used as methods of diagnosis, as described *supra.*

Kits for diagnostic use are also disclosed herein, and are comprised, within one or more containers, of an anti-RB•RB-IP complex antibody or an anti-CD24*, anti-Set*, anti-GluT1*, or anti-115392* antibody, and, optionally, a labeled binding partner to the antibody. Alternatively, the anti-RB•RB-IP complex antibody, or anti-CD24*, anti-Set*, anti-GluT1*, or anti-115392* antibody, may be labeled with a detectable marker (*e.g.*, a chemiluminescent, enzymatic,

fluorescent, or radioactive moiety) Additionally, a kit is disclosed which comprises, in one or more containers, a nucleic acid probe capable of hybridizing to *RB* and/or an *RB-IP* mRNA. In a specific embodiment of the present invention, a kit can comprise, in one or more containers, a pair of primers (*e.g.*, each in the size range of 6-30 nucleotides) that are capable of serving as primers in an amplification-based reaction, including, but not limited to: (*i*) polymerase chain reaction (PCR; see *e.g.*, Innis, *et al.*, 1990. *PCR Protocols,* Academic Press, Inc., San Diego, CA); (*ii*) ligase chain reaction; (*iii*) use of Qß replicase; (*iv*) cyclic probe reaction or various other methods known within the art, under such appropriate reaction conditions that at least a portion of an *RB* and/or an *RB-IP*, nucleic acid undergoes amplification. A kit may, optionally, further comprise (in one or more containers) a predetermined amount of a purified RB•RB-IP complex, RB and/or an RB-IP, or nucleic acids thereof , for use as, for example, a standard or control.

(V) Therapeutic Uses of RB•RB-IP COMPLEXES and CD24*, Set*, GluT1*, AND 115392*

The present invention discloses methodologies for the treatment or prevention of various diseases and disorders by administration of a therapeutic compound (hereinafter designated "Therapeutic"). Such Therapeutics may include, but are not limited to: (*i*) RB•RB-IP complexes, (*ii*) RB; (*iii*) the individual RB-IP proteins and analogs and derivatives (including fragments) of the foregoing; (*iv*) antibodies specific thereto; (*v*) nucleic acids encoding RB and/or an RB-IP, and analogs or derivatives thereof; (*vi*) RB and/or *RB-IP* antisense nucleic acids and (*vii*) RB•RB-IP complex and CD24*, Set*, GluT1*, and 115392* modulators (*e.g.*, inhibitors, agonists and antagonists).

As previously discussed *supra,* RB and several of its binding partners (IPs), as identified herein (*e.g.*, RN-tre, CAS, IK, LDH-B, Nlk1(Nek 2), CD24*, and Set*) have been implicated in playing a significant role in disorders of cell cycle progression, cell differentiation, and transcriptional control, including cancer and tumorigenesis and tumor progression. Disorders of neurodegeneration resulting from altered cellular apoptosis, differentiation, and DNA repair can likewise involve these same proteins, as well as Nlk1(Nek 2) cyclophilin A, Zap 3, and GluT1*. In accord, a wide-range of diseases affected by intracellular signal transduction and transcriptional regulation may be treated or prevented by administration of a Therapeutic which modulates (*i.e.*, inhibits, antagonizes, enhances or promotes) RB•RB-IP complex activity, or CD24*, Set*, GluT1*, or 115392* activity. Both RB and cyclophilin A have been implicated in the cellular response to viral infection.. In accord, all of these aforementioned disorders may thus be treated or prevented by administration of a Therapeutic that modulates (*i.e.*, inhibits,

antagonizes, enhances or promotes) RB•RB-IP complex activity, or CD24*, Set*, GluT1*, or 115392* activity.

Diseases or disorders associated with aberrant levels of RB•RB-IP complex levels or activity, or aberrant levels of CD24*, Set*, GluT1*, or 115392*, may be treated by administration of a Therapeutic that modulates RB•RB-IP complex formation or activity, or CD24*, Set*, GluT1*, or 115392* activity. In a specific embodiment of the present invention, the activity or level of RB is modulated by administration of an RB-IP. In another specific embodiment, the activity or level of an RB-IP is modulated by administration of RB itself.

Diseases and disorders characterized by increased (relative to a subject not suffering from the disease or disorder) RB•RB-IP levels or activity, or increased CD24*, Set*, GluT1*, or 115392* levels or activity, may be treated with Therapeutics that antagonize (*i.e.*, reduce or inhibit) RB•RB-IP complex formation or activity, or CD24*, Set*, GluT1*, and 115392* levels or activity. Therapeutics which may be used include, but are not limited to: (*i*) RB or an RB-IP, or analogs, derivatives or fragments thereof; (*ii*) anti-RB•RB-IP complex antibodies, and anti-CD24*, anti-Set*, anti-GluT1*, and anti-115392* antibodies, fragments and derivatives thereof containing the binding region thereof; (*iii*) nucleic acids encoding RB or an RB-IP; (*iv*) concurrent administration of *RB* and *RB-IP* antisense nucleic acids, or *CD24*, Set*, *GluT1**, or *115392** antisense nucleic acids; and/or (*v*) RB and/or RB-IP, or CD24*, Set*, GluT1*, or 115392* nucleic acids that are dysfunctional (*e.g.*, due to a heterologous (non-RB and/or non-RB-IP, or non-CD24*, non-Set*, non-GluT1*, or non-115392*) insertion within the coding sequences of the RB or RB-IP coding sequences) that are used to "knockout" endogenous RB and/or RB-IP function by homologous recombination (see *e.g.*, Capecchi, 1989. *Science* 244: 1288-1292).

In a specific embodiment of the present invention, a nucleic acid containing a portion of an *RB* and/or an *RB-IP* gene in which the *RB* and/or *RB-IP* sequences flank (*i.e.*, are both 5-' and 3'- to) a different gene sequence, is used, as an RB and/or an RB-IP antagonist, so as to promote *RB* and/or *RB-IP* inactivation by homologous recombination. See *e.g.*, Koller and Smithies, 1989. *Proc. Natl. Acad. Sci. USA* 86:8932-8935; Zijlstra, *et al.*, 1989. *Nature* 342:435-438. Additionally, mutants or derivatives of a first RB-IP protein that possess greater affinity for RB than that possessed by a second RB-IP, may be administered to compete with said second RB-IP protein for RB binding, thereby reducing the levels of RB complexes containing the second RB-IP. Other Therapeutics which function to inhibit RB•RB-IP complex or CD24*, Set*, GluT1*, or 115392* function may be identified by use of various *in vitro* assays well-known within the art (*e.g.*, based upon their ability to inhibit RB•RB-IP binding).

In other specific embodiments of the present invention, Therapeutics which function as antagonists in RB•RB-IP complex formation or activity, or CD24*, Set*, GluT1*, or 115392* activity, are administered therapeutically or prophylactically in the following situations: (*i*) in diseases or disorders involving an increased (relative to normal or desired) level of RB•RB-IP complex, or CD24*, Set*, GluT1*, and 115392* proteins, for example, in patients where RB•RB-IP complexes or CD24*, Set*, GluT1*, or 115392* proteins are overactive or over-expressed;or (*ii*) in diseases or disorders wherein *in vitro* and/or *in vivo* assays indicate the utility of RB•RB-IP complex, or CD24*, Set*, GluT1*, or 115392*, antagonist administration. Increased levels of RB•RB-IP complexes, or CD24*, Set*, GluT1*, or 115392* protein, may be readily detected by quantifying protein and/or RNA (*e.g.*, obtained from a patient tissue sample from biopsy tissue or the like) and performing an *in vitro* assay for RNA or protein levels, structure and/or activity of the expressed RB•RB-IP complex (or the *RB* and *RB-IP* mRNA), or CD24*, Set*, GluT1*, or 115392* protein or mRNA. Many methodologies standard within the art may thus be employed, including, but not limited to: immunoassays to detect and/or visualize RB•RB-IP complexes, or CD24*, Set*, GluT1*, or 115392* protein (*e.g.*, Western blot, immunoprecipitation followed by sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis, immunocytochemistry, etc.) and/or hybridization assays to detect concurrent expression of *RB* and an *RB-IP*, or individual *CD24*, Set*, GluT1*, and *115392** mRNA.

A specific embodiment of the present invention discloses methodologies of reducing RB•RB-IP complex expression (*i.e.*, the expression of the two components of the RB•RB-IP complex and/or formation of the complex), or CD24*, Set*, GluT1*, or 115392* expression, by targeting mRNAs that express the protein moieties. RNA-based therapeutics currently fall within three classes: (*i*) antisense species; (*ii*) ribozymes or (*iii*) RNA aptamers. See *e.g.*, Good, *et al.*, 1997. *Gene Therapy* 4:45-54.

Antisense oligonucleotides have been the most widely utilized methodology. By way of example, but not of limitation, antisense oligonucleotide methodology to reduce RB•RB-IP complex formation will be discussed fully *infra*. Ribozyme therapy involves the administration, induced expression, etc., of small RNA molecules which possess enzymatic activity to cleave, bind, or otherwise inactivate specific RNA species so as to reduce or eliminate expression of particular proteins. See *e.g.*, Grassi and Marini, 1996. *Annals of Medicine* 28:499-510; Gibson, 1996. *Cancer and Metastasis Rev.* 15:287-299. Currently, the design of "hairpin" and "hammerhead" RNA ribozymes is necessary to specifically target a particular mRNA, such as the mRNA encoding RB. RNA aptamers are specific RNA ligands for proteins (*e.g.*, such as for Tat and Rev RNA) which can specifically inhibit their translation. See *e.g.*, Good, *et al.*, 1997. *Gene*

Therapy 4:45-54. Aptamers specific for RB or an RB-IP may be identified by many methods well-known within the art (*e.g.*, protein-protein interaction assay).

In another embodiment of the present invention, the activity or level of RB is reduced by the administration of an RB-IP, or a nucleic acid that encodes an RB-IP, or antibody that immunospecifically-binds to an RB-IP, or a fragment or a derivative of the antibody containing the binding domain thereof. Additionally, the level or activity of an RB-IP may be reduced by administration of an *RB* or an *RB* nucleic acid, or an antibody that immunospecifically-binds RB, or a fragment or derivative of the antibody containing the binding domain thereof.

In yet another aspect of the present invention, diseases or disorders associated with increased levels of RB or a particular RB-IP may be treated or prevented by administration of a Therapeutic that increases RB•RB-IP complex formation, if the complex formation acts to reduce or inactivate RB or the particular RB-IP through the RB•RB-IP complex formation. Such diseases or disorders can be treated or prevented by administration of one member of the RB•RB-IP complex, including mutants of a member of the complex that possess increased affinity for the other member of the RB•RB-IP complex (so as to cause increased complex formation), administration of antibodies or other molecules that stabilize the RB•RB-IP complex, and the like.

Diseases and disorders associated with under-expression of an RB•RB-IP complex, or RB, or a particular RB-IP, are treated or prevented by administration of a Therapeutic that promotes (*i.e.*, increases or supplies) RB•RB-IP complexes or function. Examples of such a Therapeutic include, but are not limited to: (*i*) RB•RB-IP complexes and derivatives, analogs and fragments thereof that are functionally active (*e.g.*, active to form RB•RB-IP complexes); (*ii*) non-complexed RB and RB-IP proteins, and derivatives, analogs, and fragments thereof and/or (*iii*) nucleic acids encoding the members of an RB•RB-IP complex, or functionally active derivatives or fragments thereof (*e.g.*, for use in gene therapy). In a specific embodiment of the present invention are derivatives, homologs or fragments of RB and/or an RB-IP that increase and/or stabilize RB•RB-IP complex formation.

In specific embodiments, Therapeutics which promote RB•RB-IP complex function, or CD24*, Set*, GluT1*, and 115392* function, are administered therapeutically (including prophylactically) in the following clinical situations: (*i*) in diseases or disorders involving an absence or decreased (relative to normal or desired) level of RB•RB-IP complex, or CD24*, Set*, GluT1*, or 115392*, proteins, for example, in patients where RB•RB-IP complexes (or the individual components necessary to form the complexes), or CD24*, Set*, GluT1*, or 115392* is lacking, genetically defective, biologically inactive or underactive, or under-expressed or (*ii*) in

diseases or disorders wherein *in vitro* (or *in vivo*) assays indicate the utility of RB•RB-IP complex, or CD24*, Set*, GluT1*, or 115392* agonist administration. The absence or decreased level of RB•RB-IP complex, or CD24*, Set*, GluT1*, or 115392* protein or function, can be readily detected by obtaining a patient tissue sample (*e.g.*, from biopsy tissue) and performing an *in vitro* assay for RNA protein levels, activity of the expressed RB•RB-IP complex (or for the concurrent expression of mRNA encoding the two components of the RB•RB-IP complex), or CD24*, Set*, GluT1*, and 115392* RNA, protein or activity. Many methods standard within the art may be thus employed, including, but not limited to: immunoassays to detect and/or visualize RB•RB-IP complexes (or the individual components said complexes), or CD24*, Set*, GluT1*, or 115392* protein and/or hybridization assays to detect expression of the mRNA encoding the individual protein components of the RB•RB-IP complexes by detecting and/or visualizing *RB* and an *RB-IP* mRNA by use of Northern blot assays, dot blots, *in situ* hybridization, and the like.

In another specific embodiment of the present invention, the activity or level of RB is increased by administration of an RB-IP, or derivative or analog thereof, a nucleic acid encoding an RB-IP, or an antibody that immunospecifically-binds an RB-IP, or a fragment or derivative of the antibody contains the binding domain thereof. In yet another specific embodiment, the activity or levels of an RB-IP are increased by administration of RB, or derivative or analog thereof, a nucleic acid encoding RB, or an antibody that immunospecifically-binds RB, or a fragment or derivative of the antibody contains the binding domain thereof.

Generally, administration of products of a species origin or species reactivity (in the case of antibodies) that is the same species as that of the patient is preferred. Thus, in a preferred embodiment, a human RB•RB-IP complex, or CD24*, Set*, GluT1*, and 115392* protein, or derivative or analog thereof, nucleic acids encoding the members of the human RB•RB-IP complex, or human CD24*, human Set*, human GluT1*, or human 115392*, or derivative or analog thereof, or an antibody to a human RB•RB-IP complex, or CD24*, Set*, GluT1*, and 115392*, or derivative thereof, is therapeutically or prophylactically administered to a human patient.

In a preferred embodiment, suitable *in vitro* or *in vivo* assays are utilized to determine the effect of a specific Therapeutic, and whether its administration is indicated for treatment of the affected tissue. In various specific embodiments, *in vitro* assays may be carried out with representative cells or cell types involved in a patient's disorder to determine if a Therapeutic has a desired effect upon such cell types. Prior to utilization in humans, Therapeutic compounds may be tested in suitable *in vivo* animal model systems.

(a) Malignancies

As previously discussed, components of RB•RB-IP complexes (*i.e.*, RB, RN-tre, CAS, IK, LDH-B, Nlk1(Nek 2), Zap 3, CD24*, and Set*) have been implicated in regulation of cell proliferation. Accordingly, Therapeutics of the present invention may be useful in treating or preventing diseases or disorders associated with cell hyperproliferation or loss of control of cell proliferation, particularly cancers, malignancies and tumors. The Therapeutics of the present invention may be assayed by any method known within the art for efficacy in treating or preventing malignancies and related disorders. Such assays include *in vitro* assays using transformed cells or cells derived from the tumor of a patient, or *in vivo* assays using animal models of cancer or malignancies, or the like. By way of example but not of limitation, potentially effective Therapeutics may function to inhibit proliferation of tumors or transformed cells in culture, cause regression of tumors in animal models in comparison to controls, or other similar physiological response.

Accordingly, once a malignancy or cancer has been shown to be amenable to treatment by modulating (*i.e.*, inhibiting, antagonizing, enhancing or agonizing) RB•RB-IP complex activity, or CD24*, Set*, GluT1*, and 115392* activity, that cancer or malignancy can be treated or prevented by administration of a Therapeutic that modulates RB•RB-IP complex formation and function, or CD24*, Set*, GluT1*, and 115393* function, including supplying RB•RB-IP complexes and the individual binding partners of an RB•RB-IP complex. For a review of such cancers and malignancies see *e.g.*, Fishman, *et al.*, 1985. Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia.

Additionally, in specific embodiments of the present invention, malignancy or dysproliferative changes (such as metaplasias and dysplasias), or hyperproliferative disorders, are treated or prevented in, for example, the bladder, breast, colon, lung, prostate, pancreas, or uterus.

(b) Pre-Malignant Conditions

The Therapeutics of the present invention which are effective in the therapeutic or prophylactic treatment of cancer or malignancies may also be administered for the treatment of pre-malignant conditions and/or to prevent the progression of a pre-malignancy to a neoplastic or malignant state. Such prophylactic or therapeutic use is indicated in conditions known or suspected of preceding progression to neoplasia or cancer, in particular, where non-neoplastic cell growth consisting of hyperplasia, metaplasia or, most particularly, dysplasia has occurred. For a review of such abnormal cell growth see *e.g.*, Robbins & Angell, 1976. *Basic Pathology, 2nd ed.* (W.B. Saunders Co., Philadelphia, PA).

Hyperplasia is a form of controlled cell proliferation involving an increase in cell number in a tissue or organ, without significant alteration in its structure or function. For example, it has been demonstrated that endometrial hyperplasia often precedes endometrial cancer. Metaplasia is a form of controlled cell growth in which one type of mature or fully differentiated cell substitutes for another type of mature cell. Metaplasia may occur in epithelial or connective tissue cells. Dysplasia is generally considered a precursor of cancer, and is found mainly in the epithelia. Dysplasia is the most disorderly form of non-neoplastic cell growth, and involves a loss in individual cell uniformity and in the architectural orientation of cells. Dysplasia characteristically occurs where there exists chronic irritation or inflammation, and is often found in the cervix, respiratory passages, oral cavity, and gall bladder.

Alternatively, or in addition to the presence of abnormal cell growth characterized as hyperplasia, metaplasia, or dysplasia, the presence of one or more characteristics of a transformed or malignant phenotype displayed either *in vivo* or *in vitro* within a cell sample derived from a patient, is indicative of the desirability of prophylactic/therapeutic administration of a Therapeutic of the present invention which possesses the ability to modulate RB•RBIP complex activity. Characteristics of a transformed phenotype include, but are not limited to: (*i*) morphological changes; (*ii*) looser substratum attachment; (*iii*) loss of cell-to-cell contact inhibition; (*iv*) loss of anchorage dependence; (*v*) protease release; (*vi*) increased sugar transport; (*vii*) decreased serum requirement; (*viii*) expression of fetal antigens, (*ix*) disappearance of the 250 Kdal cell-surface protein, and the like. See *e.g.*, Richards, *et al.*, 1986. *Molecular Pathology* (W.B. Saunders Co., Philadelphia, PA).

In a specific embodiment of the present invention, a patient which exhibits one or more of the following predisposing factors for malignancy is treated by administration of an effective amount of a Therapeutic: (*i*) a chromosomal translocation associated with a malignancy (*e.g.*, the Philadelphia chromosome (bcr/*abl*) for chronic myelogenous leukemia and t(14;18) for follicular lymphoma, etc.); (*ii*) familial polyposis or Gardner's syndrome (possible forerunners of colon

cancer); (*iii*) monoclonal gammopathy of undetermined significance (a possible precursor of multiple myeloma) and (*iv*) a first degree kinship with persons having a cancer or pre-cancerous disease showing a Mendelian (genetic) inheritance pattern (*e.g.*, familial polyposis of the colon, Gardner's syndrome, hereditary exostosis, polyendocrine adenomatosis, medullary thyroid carcinoma with amyloid production and pheochromocytoma, Peutz-Jeghers syndrome, neurofibromatosis of Von Recklinghausen, retinoblastoma, carotid body tumor, cutaneous melanocarcinoma, intraocular melanocarcinoma, xeroderma pigmentosum, ataxia telangiectasia, Chediak-Higashi syndrome, albinism, Fanconi's aplastic anemia and Bloom's syndrome).

In another preferred embodiment, a Therapeutic of the present invention is administered to a human patient to prevent the progression to breast, colon, lung, pancreatic, uterine cancer, melanoma or sarcoma.

(c) Hyperproliferative and Dysproliferative Disorders

In one embodiment of the present invention, a Therapeutic is administered to treat or prevent hyperproliferative or benign dysproliferative disorders. Therapeutics of the present invention can be assayed by any method known in the art for efficacy in treating or preventing hyperproliferative diseases or disorders, such assays include *in vitro* cell proliferation assays, *in vitro* or *in vivo* assays using animal models of hyperproliferative diseases or disorders, or the like. Potentially effective Therapeutics include, but are not limited to, Therapeutics that reduce cell proliferation in culture or inhibit growth or cell proliferation in animal models in comparison to controls.

Accordingly, once a hyperproliferative disorder has been shown to be amenable to treatment by modulation of RB•RB-IP complex activity, or CD24*, Set*, GluT1*, or 115392* activity, that hyperproliferative disease or disorder can be treated or prevented by administration of a Therapeutic that modulates RB•RB-IP complex formation, or CD24*, Set*, GluTl* or 115392* activity (including supplying RB•RB-IP complexes and/or the individual binding partners of said complex).

Specific embodiments of the present invention are directed to treatment or prevention of cirrhosis of the liver (a condition in which scarring has overtaken normal liver regeneration processes), treatment of keloid/hypertrophic scar formation (disfiguring of the skin in which the scarring process interferes with normal renewal), psoriasis (a common skin condition characterized by excessive proliferation of the skin and delay in proper cell fate determination), benign tumors, fibrocystic conditions, and tissue hypertrophy (*e.g.*, prostatic hyperplasia).

(d) Neurodegenerative Disorders

RB and certain binding partners of RB (*e.g.*, Nlk1(Nek 2), cyclophilin A, Zap 3, and GluT1*) have been implicated in the deregulation of cellular maturation and apoptosis, which are characteristic of neurodegenerative disease (see *e.g.*, Isslebacher, *et al.*, 1997. *Principals of Internal Medicine, 13th Ed.*, McGraw Hill, New York, NY). Additionally, cyclophilin A is also implicated in mitochondrial dysfunction, again considered an underlying cause of neurodegenerative diseases and disorders. Accordingly, Therapeutics of the present invention, particularly, but not limited to those that modulate (or supply) RB•Zap 3, RB•GluT1*, RB•cyclophilin A and RB•Nlk1(Nek2) complexes maybe effective in treating or preventing neurodegenerative disease. Therapeutics of the invention which modulate RB•RB-IP complexes involved in neurodegenerative disorders can be assayed by any method known in the art for efficacy in treating or preventing such neurodegenerative diseases and disorders. Such assays include *in vitro* assays for regulated cell maturation or inhibition of apoptosis or *in vivo* assays using animal models of neurodegenerative diseases or disorders, or the like. Potentially effective Therapeutics, for example but not by way of limitation, promote regulated cell maturation and prevent cell apoptosis in culture, or reduce neurodegeneration in animal models in comparison to controls.

Once a neurodegenerative disease or disorder has been shown to be amenable to treatment by modulation of RB•RB-IP complex activity, the disease or disorder may be treated or prevented by administration of a Therapeutic that modulates RB•RB-IP complex formation.

(e) Viral Infection

As previously discussed, RB and cyclophilin A are strongly implicated in various viral infection mechanisms, including those for HIV-1.

An enormous number of human diseases result from virulent and opportunistic viral infection. Viral diseases of a wide array of tissues, including the respiratory tract, the CNS, the skin, the genitourinary tract, the eyes and ears, the immune system, the gastrointestinal tract, and the musculoskeletal system, affect a vast number of humans, of all ages.

As described *supra*, Therapeutics of the invention, particularly those that modulate (or supply) RB•RB-IP complex (particularly RB•cyclophilin A complex) activity may be effective in treating or preventing viral diseases or disorders. Therapeutics of the invention (particularly Therapeutics which modulate the levels or activity of RB•cyclophilin A) can be assayed by any method known in the art for efficacy in treating or preventing such viral diseases and disorders.

Such assays include widely used *in vitro* assays using cell culture models, and *in vivo* assays using animal models of viral diseases or disorders. See *e.g.*, McGeoch, *et al.*, 1986. *J. Gen. Virol.* 67:813-825. Potentially effective Therapeutics, by way of example but not of limitation, reduce viral responses in animal models in comparison to controls.

Accordingly, once a viral disease or disorder has been shown to be amenable to treatment by modulation of cyclophilin A complex activity, that viral disease or disorder can be treated or prevented by administration of a Therapeutic which functions to modulate RB•cyclophilin A complex formation (including supplying RB•cyclophilin A complexes).

(g) Gene Therapy

In a specific embodiment of the present invention, nucleic acids comprising a sequence encoding RB and/or an RB-IP, or CD24*, Set*, GluT1*, or 115392*, or functional derivatives thereof, are administered to modulate RB•RB-IP complexes, or CD24*, Set*, GluT1*, or 115392* function, by way of gene therapy. In additional specific embodiments, a nucleic acid or nucleic acids encoding both RB and an RB-IP (*e.g.*, RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, or 115392*), or functional derivatives thereof, are administered by way of gene therapy. Gene therapy, as utilized herein, refers to therapy performed by the administration of a nucleic acid to a subject. In this aforementioned embodiment of the present invention, the nucleic acid produces its encoded protein(s) that mediates a therapeutic effect by modulating the RB•RB-IP complex, or CD24*, Set*, GluT1*, or 115392*, function.

Any of the methods for gene therapy available within the art can be used according to the present invention. In a preferred aspect, the Therapeutic comprises a *RB* and a *RB-IP* nucleic acid, or *CD24*, Set*, *GluT1**, or *115392** nucleic acid, that is part of an expression vector that expresses the RB and RB-IP proteins, or expresses CD24*, Set*, GluT1*, or 115392*, or fragments or chimeric proteins thereof, in a suitable host. In particular, such a nucleic acid has promoter(s) operably linked to the RB and the RB-IP coding region(s), or linked to the CD24*, Set*, GluT1*, or 115392* coding region, said promoter(s) being inducible or constitutive, and, optionally, being tissue-specific. In another particular embodiment, a nucleic acid molecule is used in which the RB and RB-IP coding sequences, or CD24*, Set*, GluT1*, or 115392* coding sequences, and any other desired sequences, are flanked by regions that promote homologous recombination at a desired site in the genome, thus providing for intra-chromosomal expression of the *RB* and the *RB-IP* nucleic acids. See *e.g.*, Koller and Smithies, 1989. *Proc. Natl. Acad. Sci. USA* 86:8932-8935; Zijlstra, *et al.*, 1989. *Nature* 342:435-438.

Delivery of the Therapeutic nucleic acid into a patient may be either direct (*i.e.*, the patient is directly exposed to the nucleic acid or nucleic acid-containing vector) or indirect (*i.e.*, cells are first transformed with the nucleic acid *in vitro*, then transplanted into the patient). These two approaches are known, respectively, as *in vivo* or *ex vivo* gene therapy. In a specific embodiment of the present invention, the nucleic acid is directly administered *in vivo*, where it is expressed to produce the encoded product. This may be accomplished by any of numerous methods known in the art including, but not limited to: (*i*) constructing it as part of an appropriate nucleic acid expression vector and administering in a manner such that it becomes intracellular (*e.g.*, by infection using a defective or attenuated retroviral or other viral vector ; see U.S. Patent No. 4,980,286) or (*ii*) direct injection of naked DNA, or through the use of microparticle bombardment (*e.g.*, a "Gene Gun"; Biolistic, DuPont), or by coating it with lipids, cell-surface receptors/transfecting agents, or through encapsulation in liposomes, microparticles, or microcapsules, or by administering it in linkage to a peptide which is known to enter the nucleus, or by administering it in linkage to a ligand predisposed to receptor-mediated endocytosis (see *e.g.*, Wu & Wu, 1987. *J. Biol. Chem.* 262:4429-4432), which can be used to "target" cell types which specifically express the receptors of interest, etc.

In another specific embodiment of the present invention, a nucleic acid-ligand complex may be produced in which the ligand comprises a fusogenic viral peptide designed so as to disrupt endosomes, thus allowing the nucleic acid to avoid subsequent lysosomal degradation. In yet another specific embodiment, the nucleic acid may be targeted *in vivo* for cell-specific endocytosis and expression, by targeting a specific receptor. See *e.g.*, PCT Publications WO 92/06180; WO93/14188 and WO 93/20221. Alternatively, the nucleic acid may be introduced intracellularly and incorporated within host cell genome for expression by homologous recombination. See *e.g.*, Zijlstra, *et al.*, 1989. *Nature* 342:435-438.

In yet another specific embodiment, a viral vector that contains the *RB* and/or the *RB-IP* nucleic acids, or *CD24**, *Set**, *GluT1**, or *115392** nucleic acid, is used. For example, retroviral vectors may be employed (see *e.g.*, Miller, *et al.*, 1993. *Meth. Enzymol.* 217:581-599) which have been modified to delete those retroviral-specific sequences which are not required for packaging of the viral genome and its subsequent integration into host cell DNA. The *RB* and/or *RB-IP* (preferably both *RB* and *RB-IP*) nucleic acids, or *CD24**, *Set**, *GluT1**, or *115392** nucleic acids, to be used in gene therapy is/are cloned into the vector, which facilitates delivery of the gene into a patient. See *e.g.*, Boesen, *et al.*, 1994. *Biotherapy* 6:291-302; Kiem, *et al.*, 1994. *Blood* 83:1467-1473. Additionally, adenovirus is an especially efficacious "vehicle" for the delivery of genes to the respiratory epithelia. Other targets for adenovirus-based delivery

systems are liver, the central nervous system, endothelial cells, and muscle. Adenoviruses also possess the advantageous ability to infect non-dividing cells. For a review see *e.g.*, Kozarsky & Wilson, 1993. *Curr. Opin. Gen. Develop.* 3:499-503. Adenovirus-associated virus (AAV) has also been proposed for use in gene therapy. See *e.g.*, Walsh, *et al.*, 1993. *Proc. Soc. Exp. Biol. Med.* 204:289-300.

An additional approach to gene therapy in the practice of the present invention involves transferring a gene into cells in *in vitro* tissue culture by such methods as electroporation, lipofection, calcium phosphate-mediated transfection, or viral infection. Generally, the methodology of transfer includes the transfer of a selectable marker to the cells. The cells are then placed under selection pressure (*e.g.*, antibiotic resistance) so as facilitate the isolation of those cells which have taken up, and are expressing the transferred gene. Those cells are then delivered to a patient. In this specific embodiment, the nucleic acid is introduced into a cell prior to the *in vivo* administration of the resulting recombinant cell by any method known within the art including, but not limited to: transfection, electroporation, microinjection, infection with a viral or bacteriophage vector containing the nucleic acid sequences of interest, cell fusion, chromosome-mediated gene transfer, microcell-mediated gene transfer, spheroplast fusion, and similar methodologies which ensure that the necessary developmental and physiological functions of the recipient cells are not disrupted by the transfer. See *e.g.*, Loeffler & Behr, 1993. *Meth. Enzymol.* 217: 599-618. The technique should provide for the stable transfer of the nucleic acid to the cell, so that the nucleic acid is expressible by the cell and preferably heritable and expressible by its cell progeny.

In preferred embodiments of the present invention, the resulting recombinant cells may be delivered to a patient by various methods known within the art including, but not limited to: injection of epithelial cells (*e.g.*, subcutaneously); the application of recombinant skin cells as a skin graft onto the patient and the intravenous injection of recombinant blood cells (*e.g.*, hematopoetic stem or progenitor cells). The total amount of cells which are envisioned for use depend upon the desired effect, patient state, etc., and may be determined by one skilled within the art.

Cells into which a nucleic acid can be introduced for purposes of gene therapy encompass any desired, available cell type, and include, but are not limited to epithelial cells, endothelial cells, keratinocytes, fibroblasts, muscle cells, hepatocytes and blood cells (*e.g.*, T-lymphocytes, B-lymphocytes, monocytes, macrophages, neutrophils, eosinophils, megakaryocytes, granulocytes and hematopoetic stem or progenitor cells obtained from bone marrow, umbilical

cord blood, peripheral blood, fetal liver, etc.). In a preferred embodiment of the present invention, the cell utilized for gene therapy may be autologous to the patient.

In a specific embodiment in which recombinant cells are used in gene therapy, stem or progenitor cells, which can be isolated and maintained *in vitro*, may be utilized. Such stem cells include, but are not limited to, hematopoetic stem cells (HSC), stem cells of epithelial tissues (*e.g.*, skin, lining of the gut, embryonic heart muscle cells, liver stem cells) and neural stem cells (see *e.g.*, Stemple & Anderson, 1992. *Cell* 71:973-985). With respect to hematopoetic stem cells (HSC), any technique which provides for the isolation, propagation, and maintenance *in vitro* of HSC may be used in this specific embodiment of the invention. As previously discussed, the HSCs utilized for gene therapy are, preferably, autologous to the patient. Hence, non-autologous HSCs are, preferably, utilized in conjunction with a method of suppressing transplantation immune reactions of the future host/patient. See *e.g.*, Kodo, *et al.*, 1984. *J. Clin. Invest.* 73:1377-1384. In another preferred embodiment of the present invention, HSCs may be highly enriched (or produced in a substantially-pure form), by any techniques known within the art, prior to administration to the patient. See *e.g.*, Witlock & Witte, 1982. *Proc. Natl. Acad. Sci. USA* 79:3608-3612.

In another specific embodiment, the nucleic acid to be introduced for purposes of gene therapy comprises an inducible promoter operably linked to the coding region, such that expression of the nucleic acid is controllable by controlling the presence or absence of the appropriate inducer of transcription.

Additional methods may be adapted for use to deliver a nucleic acid encoding the RB and/or RB-IP proteins, or functional derivatives thereof, and are within the scope of the present invention.

(h) Use of Antisense Oligonucleotides for the Suppression of RB•RB-IP Complexes and CD24*, Set*, GluT1*, OR 115392*

In a specific embodiment of the present invention, RB•RB-IP complex or CD24*, Set*, GluT1*, or 115392* protein formation and function may be inhibited by the use of anti-sense nucleic acids for the RB protein and/or RB-IPS, and is preferably comprised of both the RB protein and RB-IP. In addition, the present invention discloses the therapeutic or prophylactic use of nucleic acids (of at least six nucleotides in length) which are anti-sense to a genomic sequence (gene) or cDNA encoding the RB protein and/or RB-IPS, or portions thereof. Such anti-sense nucleic acids have utility as Therapeutics which inhibit RB•RB-IP complex or CD24*,

Set*, GluT1*, or 115392* protein formation or activity, and may be utilized in a therapeutic or prophylactic manner.

Another specific embodiment of the present invention discloses methodologies for the inhibition of the expression of the RB protein and RB-IP nucleic acid sequences, within a prokaryotic or eukaryotic cell, which is comprised of providing the cell with an therapeutically-effective amount of an anti-sense nucleic acid of the RB protein and RB-IP or derivatives thereof.

The anti-sense nucleic acids of the present invention may be oligonucleotides which may either be directly administered to a cell or which may be produced *in vivo* by transcription of the exogenous, introduced sequences. In addition, the anti-sense nucleic acid may be complementary to either a coding (*i.e.*, exonic) and/or non-coding (*i.e.*, intronic) region of the RB protein or RB-IP mRNAs. The RB protein and RB-IP anti-sense nucleic acids are, at least, six nucleotides in length and are, preferably, oligonucleotides ranging from 6-200 nucleotides in length. In specific embodiments, the anti-sense oligonucleotide is at least 10 nucleotides, at least 15 nucleotides, at least 100 nucleotides, or at least 200 nucleotides. The anti-sense oligonucleotides may be DNA or RNA (or chimeric mixtures, derivatives or modified versions thereof), may be either single-stranded or double-stranded and may be modified at a base, sugar or phosphate backbone moiety.

In addition, the anti-sense oligonucleotide of the present invention may include other associated functional groups, such as peptides, moieties which facilitate the transport of the oligonucleotide across the cell membrane, a hybridization-triggered cross-linking agent, a hybridization-triggered cleavage-agent, and the like. See *e.g.*, Letsinger, *et al.*, 1989. *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; PCT Publication No. WO 88/09810. In a specific embodiment, the RB protein and RB-IP antisense oligonucleotides comprise catalytic RNAs or ribozymes. See, *e.g.*, Sarver, *et al.*, 1990. *Science* 247:1222-1225.

The anti-sense oligonucleotides of the present invention may be synthesized by standard methodologies known within the art including, but not limited to: (*i*) automated phosphorothioate-mediated oligonucleotide synthesis (see *e.g.*, Stein, *et al.*, 1988. *Nuc. Acids Res.* 16:3209) or (*ii*) methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (see *e.g.*, Sarin, *et al.*, 1988. *Proc. Natl. Acad. Sci. U.S.A.* 85:7448-7451).

In an alternative embodiment, the RB protein and RB-IP antisense nucleic acids are produced intracellularly by transcription of an exogenous sequence. For example, a vector may be produced which (upon being exocytosed by the cell) is transcribed *in vivo*, thus producing an antisense nucleic acid (RNA) species. The aforementioned vector may either remain episomal or become chromosomally-integrated, so long as it can be transcribed to produce the desired

antisense RNA. The vectors utilized in the practice of the present invention may be derived from bacterial, viral, yeast or other sources known within the art, which are utilized for replication and expression in mammalian cells. Expression of the sequences encoding the RB protein and RB-IP antisense RNAs may be facilitated by any promoter known within the art to function in mammalian, preferably, human cells. Such promoters may be inducible or constitutive and include, but are not limited to: (*i*) the SV40 early promoter region; (*ii*) the promoter contained in the 3'-terminus long terminal repeat of Rous sarcoma virus (RSV); (*iii*) the Herpesvirus thymidine kinase promoter and (*iv*) the regulatory sequences of the metallothionein gene.

The RB protein and RB-IP antisense nucleic acids may be utilized prophylactically or therapeutically in the treatment or prevention of disorders of a cell type which expresses (or preferably over-expresses) the RB protein•RB-IP complex. Cell types which express or over-express the RB protein and RB-IP RNA may be identified by various methods known within the art including, but are not limited to, hybridization with RB protein- and RB-IP-specific nucleic acids (*e.g.*, by Northern hybridization, dot blot hybridization, *in situ* hybridization) or by observing the ability of RNA from the specific cell type to be translated *in vitro* into the RB protein and the RB-IP by immunohistochemistry. In a preferred aspect, primary tissue from a patient may be assayed for the RB protein and/or RB-IP expression prior to actual treatment by, for example, immunocytochemistry or *in situ* hybridization.

Pharmaceutical compositions of the present invention, comprising an effective amount of a RB protein and RB-IP antisense nucleic acid contained within a pharmaceutically-acceptable carrier may be administered to a patient having a disease or disorder which is of a type that expresses or over-expresses RB protein•RB-IP complex RNA or protein. The amount of RB protein and/or RB-IP antisense nucleic acid which will be effective in the treatment of a particular disorder or condition will be dependant upon the nature of the disorder or condition, and may be determined by standard clinical techniques. Where possible, it is desirable to determine the antisense cytotoxicity *in vitro*, and then in useful animal model systems prior to testing and use in humans. In a specific embodiment, pharmaceutical compositions comprising RB protein and RB-IP antisense nucleic acids may be administered via liposomes, microparticles, or microcapsules. See *e.g.*, Leonetti, *et al.*, 1990. *Proc. Natl. Acad. Sci. U.S.A.* 87:2448-2451.

(VI) Assays of RB•RB-IP Complexes and CD24*, Set*, GluT1*, and 115392* Proteins

The functional activity of RB•RB-IP complexes, and CD24*, Set*, GluT1* and 115392* proteins, and derivatives, fragments and analogs thereof, can be assayed by various methods.

Potential modulators (*e.g.*, inhibitors, agonists and antagonists) of RB•RB-IP complex activity, or CD24*, Set*, GluT1*, or 115392* activity, *e.g.*, anti-RB•RB-IP, anti-CD24*, anti-Set*, anti-GluT1*, and anti-115392* antibodies, and *RB* and *RB-IP* antisense nucleic acids, can be assayed for the ability to modulate RB•RB-IP complex formation and/or activity, and CD24*, Set*, GluT1*, or 115392* activity.

(*i*) Immunoassays

For example, in one embodiment of the present invention, where one is assaying for (i) the ability to bind or compete with native, wild-type RB•RB-IP complexes, or CD24*, Set*, GluT1*, or 115392*, for (ii) the ability to bind to anti-RB•RB-IP antibodies, or anti-CD24*, anti-Set*, anti-GluT1*, or anti-115392* antibodies, various immunoassays known within the art can be used. These immunoassays include, but not limited to, competitive and non-competitive assay systems using techniques such as radioimmunoassays (RIA), enzyme linked immunosorbent assay(ELISA), "sandwich" immunoassays, immunoradiometric assays, gel diffusion precipitin reactions, immunodiffusion assays, *in situ* immunoassays, Western blots, precipitation reactions, agglutination assays, complement fixation assays, immunofluorescence assays, protein A assays, immunoelectrophoresis assays, and the like. In a specific embodiment, antibody binding is detected by detecting a label on the primary antibody. In another specific embodiment, the primary antibody is detected by detecting binding of a secondary antibody or reagent to the primary antibody. In yet another embodiment, the secondary antibody is labeled. Many means are known in the art for detecting binding in an immunoassay and are all intended to be included within the scope of the present invention.

(VII) Gene Expression Assays

The expression of the *RB* and *RB-IP* genes (both endogenous genes and those expressed from cloned DNA containing these genes) can be detected using techniques known within the art, including, but not limited to Southern hybridization; Northern hybridization; restriction endonuclease mapping; DNA sequence analysis; polymerase chain reaction amplification (PCR; see *e.g.*, U.S. Patent Nos. 4,683,202, 4,683,195, and 4,889,818; Gyllenstein, *et al.*, 1988. *Proc. Natl. Acad. Sci. USA* 85:7652-7657; Loh, *et al.*, 1989. *Science* 243:217-220) or RNase protection with probes specific for *RB* or *RB-IP* genes, in various cell types.

In one embodiment, Southern hybridization can be used to detect genetic linkage of *RB* or *RB-IP* gene mutations to physiological or pathological states. Various cell types, at various stages of development, can be characterized for their expression of RB and a RB-IP (particularly expression of RB and an RB-IP at the same time and in the same cells), or CD24*, Set*, GluT1*, or 115392* expression. The stringency of the hybridization conditions for northern or Southern blot analysis can be manipulated to ensure detection of nucleic acids with the desired degree of relatedness to the specific probes used. Modifications to these methods and other methods commonly known in the art can be used.

(VIII) Binding Assays

Derivatives (*e.g.*, fragments) and analogs of RB-IPs can be assayed for binding to RB by any method known in the art, for example (i) the modified, improved yeast two hybrid assay system (described *infra*), (ii) immunoprecipitation with an antibody that binds to RB in a complex followed by analysis by size fractionation of the immunoprecipitated proteins (*e.g.*, by denaturing or non-denaturing polyacrylamide gel electrophoresis), (iii) Western analysis, (iv) non-denaturing gel electrophoresis, and the like.

(IX) Assays for Biological Activity

A specific embodiment of the present invention provides a methodology for the screening of a derivative or analog of RB for biological activity comprising contacting said derivative or analog of RB with a RB-IP, and detecting the formation of a complex between said derivative or analog of RB and said protein; wherein detecting formation of said complex indicates that said derivative or analog of RB has biological (*e.g.*, binding) activity. Additionally, another embodiment of the invention relates to a method for screening a derivative or analog of a RB-IP protein for biological activity comprising contacting said derivative or analog of said protein with RB; and detecting the formation of a complex between said derivative or analog of said protein and RB; wherein detecting the formation of said complex indicates that said derivative or analog of said protein has biological activity.

(X) Modulation of RB•RB-IP Complex Activity

The present invention discloses methodologies relating to the modulation of the activity of a protein that possesses the ability to participate in a RB•RB-IP by administration of a binding partner of that protein, or derivative or analog thereof. RB, and derivatives and analogs thereof, can be assayed for the ability to modulate the activity or level of an RB-IP by contacting a cell or administering an animal expressing an *RB-IP* gene with an RB protein, or a nucleic acid encoding an RB protein, or an antibody that immunospecifically-binds the RB protein, or a fragment or derivative of said antibody containing the binding domain thereof, and measuring a change in RB-IP levels or activity, wherein a change in RB-IP levels or activity indicates that RB can modulate RB-IP levels or activity. Alternatively, an RB-IP can be assayed for the ability to modulate the activity or levels of an RB protein by contacting a cell or administering an animal expressing a gene encoding said protein with RB, or a nucleic acid encoding RB, or an antibody that immunospecifically-binds RB, or a fragment or derivative of said antibody containing the

binding domain thereof, wherein a change in RB levels or activity indicates that the RB-IP can modulate RB levels or activity.

(XI) RB-related Treatment Assays

(*i*) Tumorigenesis

RB, and several of the identified binding partners of RB (RB-IPs), *e.g.*, RN-tre, CAS, IK, Nlk1(Nek 2), LDH-B, CD24*, and Set*, play roles in the control of cell proliferation and, therefore, cell-transformation and tumorigenesis. The present invention discloses methodologies for screening RB•RB-IP complex and and CD24* and Set* proteins, (and derivatives, fragments, analogs and homologs, thereof) for the ability to alter cell proliferation, cell transformation and/or tumorigenesis *in vitro* and *in vivo*. For example, but not by way of limitation, cell proliferation may be assayed by measuring ^{3}H-thymidine incorporation, by direct cell count, by detecting changes in transcriptional activity of known genes such as proto-oncogenes (*e.g.*, c-*fos*, c-*myc*) cell-cycle markers, and the like.

The RB•RB-IP complex and and CD24* and Set* proteins (and derivatives, fragments, analogs and homologs, thereof) may also be screened for activity in inducing or inhibiting cell transformation (or the progression to malignant phenotype) *in vitro*. The proteins and protein complex of the present invention may be screened by contacting either cells with a normal phenotype (for assaying for cell transformation) or a transformed cell phenotype (for assaying for inhibition of cell transformation) with the protein or protein complex of the present invention and examining the cells for acquisition or loss of characteristics associated with a transformed phenotype (a set of *in vitro* characteristics associated with a tumorigenic ability *in vivo*) including, but not limited to: colony formation in soft agar, a more rounded cell morphology, looser substratum attachment, loss of contact inhibition, loss of anchorage dependence, release of proteases such as plasminogen activator, increased sugar transport, decreased serum requirement, expression of fetal antigens, disappearance of the 250 Kdal cell-surface protein, and the like. See *e.g.*, Luria, *et al.*, 1978. *General Virology, 3rd ed.* (John Wiley & Sons, New York, NY).

The RB•RB-IP complex and and CD24* and Set* proteins, (and derivatives, fragments, analogs and homologs, thereof) may also be screened for activity to promote or inhibit tumor formation *in vivo* in non-human test animal. A vast number of animal models of hyperproliferative disorders (*e.g.*, tumorigenesis and metastatic spread) are known within the art. See *e.g.*, Lovejoy, *et al.*, 1997. *J. Pathol.* 181:130-135. In a specific embodiment of the present invention, the proteins and protein complex may be administered to a non-human test animal

(preferably a test animal predisposed to develop a type of tumor) and the non-human test animals is subsequently examined for an increased incidence of tumor formation in comparison with controls animals which were not administered the proteins or protein complex of the present invention. Alternatively, the proteins and protein complex may be administered to non-human test animals possessing tumors (*e.g.*, animals in which tumors have been induced by introduction of malignant, neoplastic, or transformed cells or by administration of a carcinogen) and subsequently examining the tumors within the test animals for tumor regression in comparison to controls. Accordingly, once a hyperproliferative disease or disorder has been shown to be amenable to treatment by modulation of RB•RB-IP complex activity that disease or disorder may be treated or prevented by administration of a Therapeutic which modulates RB•RB-IP complex formation.

(*ii*) Neurodegenerative Diseases

As well as the complexes and proteins previously described *supra,* which are implicated in the control of cell cycle progression, RB, and the RB-IPs cyclophilin A, Zap 3, 17 beta-HSD6, and GluT1*, specifically have been implicated in cellular apoptosis. In addition, in the case of Zap 3, expression of this protein is genetically-linked to neurodegenerative disease. The RB•RB-IP complexes (particularly the RB•cyclophilin A, RB•Zap 3, RB•17 beta-HSD6, and RB•GluT1* complexes) and derivatives, analogs and fragments thereof, nucleic acids encoding the *RB* and *RB-IP* genes, anti-RB•RB-IP antibodies, and other modulators of RB•RB-IP complex activity, may be tested for activity in treating or preventing neurodegenerative disease in *in vitro* and *in vivo* assays.

In one embodiment of the present invention, a Therapeutic may be assayed for activity in treating or preventing neurodegenerative disease by contacting cultured cells which exhibit an indicator of a neurodegenerative disease. By way of example, but not of limitation, an assay of the level of over-expression of the β-A4 peptide, *in vitro* may be used to ascertain the efficacy of a specific Therapeutic by comparing the level of said indicator within the cells contacted with the Therapeutic with said level of said indicator in cells not so contacted, wherein a lower level in said contacted cells indicates that the Therapeutic has activity in treating or preventing neurodegenerative disease. Specific examples of cell culture models for neurodegenerative disease include, but are not limited to: (*i*) cultured rat endothelial cells from affected and non-affected individuals (see *e.g.*, Maneiro, *et al.*, 1997. *Methods Find. Exp. Clin. Pharmacol.* 19:5-12); (*ii*) P19 murine embryonal carcinoma cells (see *e.g.*, Hung, *et al.*, 1992. *Proc. Natl. Acad. Sci. USA* 89:9439-9443) and (*iii*) dissociated cell cultures of cholinergic neurons from the

nucleus basalis of Meynert (see *e.g.*, Nakajima, *et al.*, 1985. *Proc. Natl. Acad. Sci. USA* 82:6325-6329).

In another embodiment of the present invention, a Therapeutic may be assayed for activity in treating or preventing neurodegenerative disease by administering the Therapeutic to a test animal which exhibits symptoms of a neurodegenerative disease (*e.g.*, premature development of cognitive deficiencies in transgenic animals expressing β-APP) or that is predisposed to develop symptoms of a neurodegenerative disease; and measuring the change in said symptoms of the neurodegenerative disease following the administration of said Therapeutic; wherein a reduction in the severity of the symptoms of the neurodegenerative disease, or prevention of the symptoms of the neurodegenerative disease, indicates that the Therapeutic possesses activity in treating or preventing neurodegenerative disease. Such a test animal may be any one of a number of animal models known within the art for neurodegenerative disease. These aforementioned animal models (including those for Alzheimer's Disease and trisomy 21), accurately mimic natural human neurodegenerative diseases. See *e.g.*, Campbell, *et al.*, 1997. *Mol. Psychiatry* 2:125-129; Schultz, *et al.*, 1997. *Mol. Cell. Biochem.* 174:193-197. Examples of specific neurodegenerative animals models include, but are not limited to: (*i*) the partial trisomy 16 mouse (see *e.g.*, Holtzman, *et al.*, 1996. *Proc. Natl. Acad. Sci. USA* 93:13333-13338); (*ii*) bilateral nucleus basalis magnocellularis-lesioned rats (see *e.g.*, Popovic, *et al.*, 1996. *Int. J. Neurosci.* 86:281-299); (*iii*) the aged rat (see *e.g.*, Muir, 1997. *Pharmacol. Biochem. Behav.* 56:687-696); (*iv*) the PDAPP transgenic mouse model of Alzheimer's disease (see *e.g.*, Johnson-Wood, *et al.*, 1997. *Proc. Natl. Acad. Sci. USA* 94:1550-1555) and (*v*) experimental autoimmune dementia (see *e.g.*, Oron, *et al.*, 1997. *J. Neural Transm. Suppl.* 49:77-84).

(*iii*) Viral Infection

In addition, the interactants RB and cyclophilin A are strongly implicated in viral infection mechanisms, including that for HIV-1. As previously discussed, numerous human diseases directly result from virulent and opportunistic viral infection. See *e.g.*, Belshe, 1984. *Textbook of Human Virology*, PSG Publishing, Littleton, MA. Accordingly, RB•RB-IP complexes (particularly RB•cyclophilin A complexes), and derivatives, analogs, and fragments thereof, nucleic acids containing the *RB* and *RB-IP* genes (in particular the *cyclophilin A* gene); anti-RB•RB-IP antibodies (in particular anti-RB•cyclophilin A antibodies) and various other modulators of the RB•RB-IP complex activity, may be tested for activity in treating or preventing viral diseases in both *in vitro* and *in vivo* assays.

In one embodiment of the present invention, a Therapeutic may be assayed for activity in treating or preventing viral disease by contacting cultured cells that exhibit an indicator of a viral reaction (*e.g.*, formation of inclusion bodies) *in vitro* with the Therapeutic, and comparing the level of said indicator in the cells contacted with the Therapeutic with said level of said indicator in cells not so contacted, wherein a lower level in said contacted cells indicates that the Therapeutic has activity in treating or preventing viral disease. Cell models which may be used for such assays include, but are not limited to: (*i*) viral infection of T lymphocytes (see *e.g.*, Selin, *et al.*, 1996. *J. Exp. Med.* 183:2489-2499); (*ii*) hepatitis B infection of de-differentiated hepatoma cells (see *e.g.*, Raney, *et al.*, 1997. *J. Virol.* 71:1058-1071); (*iii*) viral infection of cultured salivary gland epithelial cells (see *e.g.*, Clark, *et al.*, 1994. *Autoimmunity* 18:7-14); (*iv*) synchronous HIV-1 infection of $CD4^+$ lymphocytic cell lines (see *e.g.*, Wainberg, *et al.*, 1997. *Virology* 233:364-373); (*v*) viral infection of respiratory epithelial cells (see *e.g.*, Stark, *et al.*, 1996. *Human Gene Ther.* 7:1669-1681) and (*vi*) amphotrophic retroviral infection of NIH-3T3 cells (see *e.g.*, Morgan, *et al.*, 1995. *J. Virol.* 69:6994-7000).

In yet another embodiment, a Therapeutic of the present invention my be assayed for activity in treating or preventing viral disease by administering said Therapeutic to a test animal having symptoms of a viral infection (*e.g.*, characteristic respiratory symptoms in animal models of virally-induced asthma), or which test animal does not exhibit a viral reaction and is subsequently challenged with an agent that elicits an viral reaction, and measuring the change in the viral reaction after the administration of said Therapeutic; wherein a reduction in said viral reaction or a prevention of said viral reaction indicates that the Therapeutic has activity in treating or preventing viral disease. Animal models which may be utilized for such assays include, but are not limited to: (*i*) respiratory viral infections in guinea pigs (see *e.g.*, Kudlacz and Knippenberg, 1995. *Inflamm. Res.* 44:105-110); (*ii*) influenza virus infection of mice (see *e.g.*, Dobbs, *et al.*, 1996. *J. Immunol.* 157:1870-1877); (*iii*) neurotrophic virus infection of mice (see *e.g.*, Barna, *et al.*, 1996. *Virology* 223:331-343); (*iv*) measles infection in hamsters (see *e.g.*, Fukuda, *et al.*, 1994. *Acta Otolaryngol. Suppl (Stockh.)* 514:111-116); (*v*) encephalomyocarditis infection of mice (see *e.g.*, Hirasawa, *et al.*, 1997. *J. Virol.* 71:4024-4031) and (*vi*) cytomegalovirus (CMV) infection of mice (see *e.g.*, Orange and Biron, 1996. *J. Immunol.* 156:1138-1142).

(XII) Assays for Protein•Protein Interactions

The present invention discloses methodologies for assaying and screening derivatives, fragments, analogs and homologs of RB-IP for binding to RB. The derivatives, fragments,

analogs and homologs of the RB-IP which interact with RB may be identified by means of a yeast two hybrid assay system (see *e.g.*, Fields & Song, 1989. *Nature* 340:245-246) or; preferably, a modification and improvement thereof, as described in U.S. Patent Applications Serial Nos. 08/663,824 (filed June 14, 1996) and 08/874,825 (filed June 13, 1997), both of which are entitled "Identification and Comparison of Protein-Protein Interactions that Occur in Populations and Identification of Inhibitors of These Interactions," to Nandabalan, *et al.*, and which are incorporated by reference herein in their entireties.

The identification of interacting proteins by the improved yeast two hybrid system is based upon the detection of the expression of a reporter gene (hereinafter "Reporter Gene"), the transcription of which is dependent upon the reconstitution of a transcriptional regulator by the interaction of two proteins, each fused to one half of the transcriptional regulator. The bait RB (or derivative, fragment, analog or homolog) and prey protein (proteins to be tested for ability to interact with the bait protein) are expressed as fusion proteins to a DNA-binding domain, and to a transcriptional regulatory domain, respectively, or *vice versa*. In a specific embodiment of the present invention, the prey population may be one or more nucleic acids encoding mutants of RB-IP (*e.g.*, as generated by site-directed mutagenesis or another method of producing mutations in a nucleotide sequence). Preferably, the prey populations are proteins encoded by DNA (*e.g.*, cDNA, genomic DNA or synthetically generated DNA). For example, the populations may be expressed from chimeric genes comprising cDNA sequences derived from a non-characterized sample of a population of cDNA from mammalian RNA. In another specific embodiment, recombinant biological libraries expressing random peptides may be used as the source of prey nucleic acids.

The present invention discloses methods for the screening for inhibitors of RB-IP. In brief, the protein-protein interaction assay may be performed as previously described herein, with the exception that it is performed in the presence of one or more candidate molecules. A resulting increase or decrease in Reporter Gene activity, in relation to that which was present when the one or more candidate molecules are absent, indicates that the candidate molecule exerts an effect on the interacting pair. In a preferred embodiment, inhibition of the protein interaction is necessary for the yeast cells to survive, for example, where a non-attenuated protein interaction causes the activation of the *URA3* gene, causing yeast to die in medium containing the chemical 5-fluoroorotic acid. See *e.g.*, Rothstein, 1983. *Meth. Enzymol.* 101:167-180.

In general, the proteins comprising the bait and prey populations are provided as fusion (chimeric) proteins, preferably by recombinant expression of a chimeric coding sequence containing each protein contiguous to a pre-selected sequence. For one population, the pre-

selected sequence is a DNA-binding domain that may be any DNA-binding domain, so long as it specifically recognizes a DNA sequence within a promoter (*e.g.*, a transcriptional activator or inhibitor). For the other population, the pre-selected sequence is an activator or inhibitor domain of a transcriptional activator or inhibitor, respectively. The regulatory domain alone (not as a fusion to a protein sequence) and the DNA-binding domain alone (not as a fusion to a protein sequence) preferably, do not detectably interact, so as to avoid false-positives in the assay. The assay system further includes a reporter gene operably linked to a promoter that contains a binding site for the DNA-binding domain of the transcriptional activator (or inhibitor). Accordingly, in the practice of the present invention, the binding of the RB fusion protein to a prey fusion protein leads to reconstitution of a transcriptional activator (or inhibitor), which concomitantly activates (or inhibits) expression of the Reporter Gene.

In a specific embodiment, the present invention discloses a methodology for detecting one or more protein-protein interactions comprising the following steps: (*i*) recombinantly-expressing the RB (or a derivative, fragment, analog or homolog thereof) in a first population of yeast cells of a first mating type and possessing a first fusion protein containing the RB sequence and a DNA-binding domain; wherein said first population of yeast cells contains a first nucleotide sequence operably-linked to a promoter which is "driven" by one or more DNA-binding sites recognized by said DNA-binding domain such that an interaction of said first fusion protein with a second fusion protein (comprising a transcriptional activation domain) results in increased transcription of said first nucleotide sequence; (*ii*) negatively selecting to eliminate those yeast cells in said first population in which said increased transcription of said first nucleotide sequence occurs in the absence of said second fusion protein; (*iii*) recombinantly expressing in a second population of yeast cells of a second mating type different from said first mating type, a plurality of said second fusion proteins; wherein said second fusion protein is comprised of a sequence of a derivative, fragment, analog or homolog of a RB-IP and an activation domain of a transcriptional activator, in which the activation domain is the same in each said second fusion protein; (*iv*) mating said first population of yeast cells with said second population of yeast cells to form a third population of diploid yeast cells, wherein said third population of diploid yeast cells contains a second nucleotide sequence operably linked to a promoter "driven" by a DNA-binding site recognized by said DNA-binding domain such that an interaction of a first fusion protein with a second fusion protein results in increased transcription of said second nucleotide sequence, in which the first and second nucleotide sequences can be the same or different and (*v*) detecting said increased transcription of said first and/or second

nucleotide sequence, thereby detecting an interaction between a first fusion protein and a second fusion protein.

In a preferred embodiment, the bait (a RB sequence) and the prey (a library of chimeric genes) are combined by mating the two yeast strains on solid media for a period of approximately 6-8 hours. In a less preferred embodiment, the mating is performed in liquid media. The resulting diploids contain both types of chimeric genes (*i.e.*, the DNA-binding domain fusion and the activation domain fusion). After an interactive population is obtained, the DNA sequences encoding the pairs of interactive proteins are isolated by a method wherein either the DNA-binding domain hybrids or the activation domain hybrids are amplified, in separate reactions. Preferably, the amplification is carried out by polymerase chain reaction (PCR; see *e.g.*, Innis, *et al.*, 1990. *PCR Protocols* (Academic Press, Inc., San Diego, CA)) utilizing pairs of oligonucleotide primers specific for either the DNA-binding domain hybrids or the activation domain hybrids. The PCR amplification reaction may also be performed on pooled cells expressing interacting protein pairs, preferably pooled arrays of interactants. Other amplification methods known within the art may also be used including, but not limited to, ligase chain reaction; Qβ-replicase or the like. See *e.g.*, Kricka, *et al.*, 1995. *Molecular Probing, Blotting, and Sequencing* (Academic Press, New York, NY).

In an additional embodiment of the present invention, the plasmids encoding the DNA-binding domain hybrid and the activation domain hybrid proteins may also be isolated and cloned by any of the methods well-known within the art. For example, but not by way of limitation, if a shuttle (yeast to *E. coli*) vector is used to express the fusion proteins, the genes may be subsequently recovered by transforming the yeast DNA into *E. coli* and recovering the plasmids from the bacteria. See *e.g.*, Hoffman ,*et al.*, 1987. *Gene* 57:267-272.

(XIII) Pharmaceutical Compositions and Therapeutic/Prophylactic Administration

The invention present discloses methods of treatment and prophylaxis by the administration to a subject of an pharmaceutically-effective amount of a Therapeutic of the invention. In a preferred embodiment, the Therapeutic is substantially purified and the subject is a mammal, and most preferably, human.

Formulations and methods of administration that can be employed when the Therapeutic comprises a nucleic acid are described in Sections 6(*i*) and 6(*ii*), *supra*. Various delivery systems are known and can be used to administer a Therapeutic of the present invention including, but not limited to: (*i*) encapsulation in liposomes, microparticles, microcapsules; (*ii*) recombinant cells capable of expressing the Therapeutic; (*iii*) receptor-mediated endocytosis (see, *e.g.*, Wu & Wu,

1987. *J. Biol. Chem.* 262:4429-4432); (*iv*) construction of a Therapeutic nucleic acid as part of a retroviral or other vector, and the like.

Methods of administration include, but are not limited to, intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, epidural, and oral routes. The Therapeutics of the present invention may be administered by any convenient route, for example by infusion or bolus injection, by absorption through epithelial or mucocutaneous linings (*e.g.*, oral mucosa, rectal and intestinal mucosa, etc.) and may be administered together with other biologically-active agents. Administration can be systemic or local. In addition, it may be advantageous to administer the Therapeutic into the central nervous system by any suitable route, including intraventricular and intrathecal injection. Intraventricular injection may be facilitated by an intraventricular catheter attached to a reservoir (*e.g.*, an Ommaya reservoir). Pulmonary administration may also be employed by use of an inhaler or nebulizer, and formulation with an aerosolizing agent. It may also be desirable to administer the Therapeutic locally to the area in need of treatment; this may be achieved by, for example, and not by way of limitation, local infusion during surgery, topical application, by injection, by means of a catheter, by means of a suppository, or by means of an implant. In a specific embodiment, administration may be by direct injection at the site (or former site) of a malignant tumor or neoplastic or pre-neoplastic tissue.

In another embodiment of the present invention, the Therapeutic may be delivered in a vesicle, in particular a liposome. See *e.g.*, Langer, 1990. *Science* 249:1527-1533. In yet another embodiment, the Therapeutic can be delivered in a controlled release system including, but not limited to: a delivery pump (see *e.g.*, Saudek, *et al.*, 1989. *New Engl. J. Med.* 321:574 and a semi-permeable polymeric material (see *e.g.*, Howard, *et al.*, 1989. *J. Neurosurg.* 71:105). Additionally, the controlled release system can be placed in proximity of the therapeutic target (*e.g.*, the brain), thus requiring only a fraction of the systemic dose. See, *e.g.*, Goodson, In: *Medical Applications of Controlled Release* 1984. (CRC Press, Bocca Raton, FL).

In a specific embodiment of the present invention, where the Therapeutic is a nucleic acid encoding a protein, the Therapeutic nucleic acid may be administered *in vivo* to promote expression of its encoded protein, by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular (*e.g.*, by use of a retroviral vector, by direct injection, by use of microparticle bombardment, by coating with lipids or cell-surface receptors or transfecting agents, or by administering it in linkage to a homeobox-like peptide which is known to enter the nucleus (see *e.g.*, Joliot, *et al.*, 1991. *Proc. Natl. Acad. Sci. USA* 88:1864-1868), and the like. Alternatively, a nucleic acid Therapeutic can be introduced

intracellularly and incorporated within host cell DNA for expression, by homologous recombination.

The present invention also provides pharmaceutical compositions. Such compositions comprise a therapeutically-effective amount of a Therapeutic, and a pharmaceutically acceptable carrier. As utilized herein, the term "pharmaceutically acceptable" means approved by a regulatory agency of the Federal or a state government or listed in the U.S. Pharmacopoeia or other generally recognized pharmacopoeia for use in animals and, more particularly, in humans. The term "carrier" refers to a diluent, adjuvant, excipient, or vehicle with which the therapeutic is administered and includes, but is not limited to such sterile liquids as water and oils.

The amount of the Therapeutic of the invention which will be effective in the treatment of a particular disorder or condition will depend on the nature of the disorder or condition, and may be determined by standard clinical techniques by those of average skill within the art. In addition, *in vitro* assays may optionally be employed to help identify optimal dosage ranges. The precise dose to be employed in the formulation will also depend on the route of administration, and the overall seriousness of the disease or disorder, and should be decided according to the judgment of the practitioner and each patient's circumstances. However, suitable dosage ranges for intravenous administration of the Therapeutics of the present invention are generally about 20-500 micrograms (μg) of active compound per kilogram (Kg) body weight. Suitable dosage ranges for intranasal administration are generally about 0.01 pg/kg body weight to 1 mg/kg body weight. Effective doses may be extrapolated from dose-response curves derived from *in vitro* or animal model test systems. Suppositories generally contain active ingredient in the range of 0.5% to 10% by weight; oral formulations preferably contain 10% to 95% active ingredient.

The present invention also provides a pharmaceutical pack or kit, comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions and Therapeutics of the present invention. Optionally associated with such container(s) may be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration.

SPECIFIC EXAMPLES

(A) Identification of RB•RB-IP Complexes

A modified, improved yeast two hybrid system was used to identify protein interactions. Yeast is a eukaryote, and therefore, intermolecular protein interactions detected in this type of

system would be expected to demonstrate protein interactions that occur under physiological conditions. See *e.g.*, Chien, *et al.*, 1991. *Proc. Natl. Acad. Sci. USA* 88:9578-9581. Expression vectors were constructed to encode two hybrid proteins. For the "forward screen", one hybrid consisted of the DNA binding domain of the yeast transcriptional activator Gal4 fused to a portion of RB; whereas the other hybrid consisted of the Gal4 activator domain fused to "prey" protein sequences encoded by a mammalian cDNA library. Each of the vectors was then inserted into complementary (a and alpha) mating types of yeast using methods known within the art. See *e.g.*, Chien, *et al.*, 1991. *Proc. Natl. Acad. Sci. USA* 88:9578-9581. Mating of the yeast strains was carried out to express both vector constructs within the same yeast cells, thus allowing interaction to occur. Interaction between the bait and prey domains led to transcriptional activation of Reporter Genes containing *cis*-binding elements for Gal4. The Reporter Genes encoding the indicator protein beta-galactosidase, and metabolic markers for uracil and histidine auxotrophy, were included in specific fashion in one or the other of the yeast strains used in the mating. In this manner, yeast were selected for successful mating, expression of both fusion constructs, and expression of RB-IPs. Yeast clones which contained interacting regions were picked and grown in individual wells of microtiter plates. The plasmids containing the RB-IPs were then isolated and characterized.

The prey cDNAs were obtained from a commercial human fetal brain cDNA library of 3.5×10^6 independent isolates (Cat. #HL4029AH; Clontech, Palo Alto, CA). The library was synthesized from Xho 1-dT_{15} primed fetal brain mRNA (from a mixture of five male and female 19-22 week fetuses) that was directionally cloned into pACT2, a yeast Gal4 activation domain cloning vector including the *LEU2* gene for selection in yeast deficient in leucine biosynthesis.

The RB binding domain clone was constructed by inserting a BamH1/Sal1 fragment of *RB* from amino acid residues 301 to 928 of RB (GenBank Accession Number M28419; nucleotides 1040 to 2925 between the BamH1 and Sal1 sites of pAS-1 (Clontech). This vector is a yeast DNA-binding domain cloning vector that contains the *TRP1* gene for selection in yeast strains deficient in tryptophan biosynthesis. The bait sequence was confirmed by nucleic acid sequencing to confirm that PCR amplification reproduced an accurate copy of the RB sequence. This test determined that as predicted, the bait sequence encoded an interacting domain identical to human RB amino acid residues 301 to 928.

The bait was expressed by lithium acetate/polyethylene glycol transformation (see *e.g.*, Ito, *et al.*, 1983. *J. Bacteriol.* 153:163-168) into the yeast strain YULH (mating type a, *ura3, his3, lys2, Ade2, trp1, leu2, gal4, gal80, GAL1-URA3, GAL1-lacZ)*, while the prey sequences were expressed by transformation into the yeast strain N106' (mating type α, *ura3, his3, ade2,*

trp1, leu2, gal4, gal80, cyh^r, Lys2::GAL1$_{UAS}$-HIS3$_{TATA}$-HIS3, ura 3::GAL1$_{UAS}$-GAL$_{TATA}$-lacZ). The two transformed populations were then mated using standard methods within the art. In brief, the cells were grown until mid-to-late log phase on media that selected for the presence of the appropriate plasmids. The two mating strains, alpha and a, were then, diluted in YAPD media (see *e.g.*, Sherman, *et al.*, 1991. *Getting Started with Yeast*, Vol. 194, Academic Press, New York, NY), filtered onto nitrocellulose membranes, and incubated at 30°C for 6-8 hours. The cells were then transferred to medium selective for the desired diploids (*i.e.*, yeast harboring reporter genes for β-galactosidase, uracil auxotrophy, and histidine auxotrophy, and expression of the vectors encoding the bait and prey). The mating products were then plated on SC (synthetic complete) medium (see *e.g.*, Kaiser, 1994. *Methods in Yeast Genetics, 1994 Ed.*, Cold Spring Harbor Laboratory Press, New York, NY) lacking adenine and lysine (to select for successful mating), leucine and tryptophan (to select for expression of genes encoded by both the bait and prey plasmids), and uracil and histidine (to select for protein interactions). This medium is hereinafter designated SCS medium, for SC Selective medium.

The selected clones were tested for expression of β-galactosidase to confirm the formation of a RB•RB-IP interaction. Filter-lift β-galactosidase assays were then performed as a modified version of the protocol of Breeden and Nasmyth (1985. *Cold Spring Harbor Quant. Biol.* 50:643-650). Colonies were patched onto SCS plates, grown overnight, and replica plated onto Whatman No. 1 filters. The replica filters were subsequently assayed for β-galactosidase activity (i.e., positive colonies turned a visible blue).

The cells contained within colonies which were positive for protein interaction contained a mixture of DNA-binding and activation-domain plasmids. These cells were individually plated, and regrown as single isolates in individual wells of 96-well plates. Ten microliters of each isolate were lysed, the inserts within the pACT2 and pASS-l plasmids were amplified by polymerase chain reaction using primers specific for the flanking sequences of each vector, and approximately 200 amino-terminal bases of each insert were determined using an ABI Model 377 sequenator. Comparison to known sequences was made using the "BLAST" program publicly available through the National Center for Biotechnology Information.

Identified sequences included the RN-tre nucleic acid sequence, starting from nucleotide 2551, the CAS sequence starting from nucleotide 2263, the L6 sequence starting from nucleotide 217, the IK sequence starting from nucleotide 4, the LDH-B sequence starting from nucleotide 847, the Nlk1(Nek 2) sequence starting from nucleotide 1089, cyclophilin A sequence starting from nucleotide 3, the Zap 3 sequence starting from nucleotide 886, the CD24 sequence starting from nucleotide 689 (Figure 1, SEQ ID NO:1), the Set sequence starting from nucleotide 2076

(Figure 3, SEQ ID NO:6),, the GluT1 sequence starting from nucleotide 3408 (Figure 5, SEQ. ID. NO:9), mRNA 115392* sequence starting from nucleotide 343 (Figure 7, SEQ. ID. NO:12). and two isolates identical to the IP-30 sequence starting from nucleotides 464 and 476.

The determined nucleic acid sequences and corresponding amino acid sequences of the splice variants Set* and GluT1*, and the open reading frames encoding CD24* and 115392*, are shown in Figures 2, 4, 6, and 7, respectively. Determination of the splice variants sequences for Set*, and GluT1*, and the open reading frames (ORFs) encoding proteins CD24* and 115392*, is described in the Specific Examples Section, *infra.*

(B) Verification of the Specificity of RB•RB-IP Interactions

To test for the specificity of bait:prey interactions, two general tests were first performed. In the first instance, YULH yeast cells were created that express the individual plasmids encoding RB, RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392*. These yeast cells were plated on SCS plates, grown overnight, and examined for growth. No growth was found for RB and all RB-IP proteins, confirming that they were not "self-activating" proteins, that is, these proteins require interaction with a second protein domain for a functional activation complex.

In the second instance, plasmids containing RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392* inserts were transformed into strain N106^r (mating type alpha) and mated with yeast strain YULH (mating type a) expressing proteins other than RB. Promiscuous binders, that is, RB-IP able to bind with many other proteins in a non-specific fashion, would interact non-specifically with non-RB domains, and would be discarded as non-specific interactants. None of the interactants showed binding to proteins other than those described in the following paragraph.

To recapitulate the detected interactions, and further demonstrate their specificity, the isolated bait plasmid for RB, and separately the plasmid encoding human bait protein 1 (B1) were used to transform yeast strain YULH (mating type a). The interacting domains detected in RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, protein 2 (P2), CD24*, Set*, GluT1*, and 115392*, as well as the known interactant RbAP46 (see *e.g.*, Huang, *et al.*, 1991. *Nature* 350:160-162), were transformed into strain N106^r(mating type alpha). The transformants were regrown, and a mating performed to recapitulate the identified RB•RB-IP interactions. As shown in Figure 8, RB complexed specifically with RN-tre (Box A), CAS (Box B), IP-30 (Box C), L6 (Box D), IK (Box E), LDH-B (Box F), Nlk1(Nek 2) (Box G), cyclophilin A (Box I), Zap 3 (Box J), P2 (Box K), CD24* (Box L), Set* (Box M), GluT1* (Box N), and

115392* (Box O), as well as the known interactant RbAP46 (Box P). RB did not react non-specifically with the prey P1. As illustrated in Figure 8, the intersection of the RB row (top) with the RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, 17 beta-HSD6, CD24*, Set*, GluT1*, 115392*, and RbAP26 columns indicates growth (*i.e.*, a positive interaction), but the intersection of the RB row with the column for P1 indicates no growth (*i.e.*, no protein interaction). The interaction of B1 with all prey proteins except P1 resulted in no growth (*i.e.*, no interaction), thus revealing that the prey interactants RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek 2), cyclophilin A, Zap 3, P2, CD24*, Set*, GluT1*, 115392*, and RbAP46 were not non-specific interactors. Mating of B1 with P1 (Box Q) demonstrates that P1 is capable of forming functional interactions.

(C) Identification of the Sequences Encoding CD24*, Set*, GluT1* and 115392*

Regions within the 3' untranslated regions of the known protein cDNAs for CD24, Set, and GluT1 were identified as encoding a protein or proteins that interact with RB using the improved, modified yeast two hybrid system. The present invention discloses interacting nucleotide sequences identical to the nucleotide sequence of CD24 3' of nucleotide 689 (as depicted in Figure 1; SEQ ID NO:1); the nucleotide sequence of Set 3' of nucleotide 2175 (Figure 3; SEQ ID No:6) and the nucleotide sequence of GluT1 3' of nucleotide 3408 (as depicted in Figure 5; SEQ ID No:9). Furthermore, the human expressed clone 115392 was detected as an interacting prey, 3' of nucleotide 343 of clone 115392 sequence. This represented the first evidence that the 115392 sequence encodes a functional protein or protein domain.

For the four interactant sequences of CD24, Set, Glut, and 115392, an analysis was performed in order to ascertain whether there was any open reading frames (ORFs) within each of the three forward translation frames. The interacting protein domains were encoded by sequences of known orientation within the prey sequence vector, and thus only forward translations could explain the detected interacting species. Open reading frames (ORFs) were defined by performing a "BLAST" analysis with "ORF Finder". For CD24, only two open reading frames were detected. The first ORF encoded the known CD24 protein (indicated as A in Figure 1, from nucleotides 57-299; and as Frame A in Figure 2). The second open reading frame (indicated as C in Figure 1, from nucleotides 957-1151), included the detected interacting region for the CD24 sequence. This novel protein, called CD24*, is depicted in Figure 2, Frame B [SEQ ID NO:4 for the nucleic acid sequence and SEQ ID NO:5 for the amino acid sequence]. It should be noted that no splice variant containing the interacting domain and an upstream 5'-coding sequence could be detected, using the methods described immediately above. For the

115392 sequence, only one open reading frame, beginning at nucleotide 245 with an initiator methionine codon, and ending at nucleotide 721, the last nucleotide of a TGA translational stop codon, was found. This sequence, and the translated protein referred to as 115392*, is illustrated in Figure 7 [SEQ ID NOS:12 and 13, respectively].

No ORF of 60 amino acids or greater, beginning with an initiator methionine, and no ORF beginning from the 5'-terminus that could represent the C-terminus of a protein of 60 amino acids or longer, was detected that contains the nucleotide sequence encoding the interacting regions of either Set or GluT1. Thus, the sequences were examined to determine if they could encode splice variants of the known Set and GluT1 proteins that included the detected interacting sequences.

Determination of 5'- and 3'-splice points for protein splice variants was performed as follows. First, potential 5'-splice sites were identified in the coding sequence of the known protein. The sequence must contain an invariant GT sequence at the start of the intron. The remaining nucleotides were not invariant, but the preferred consensus sequence was AG:GTAAGT, with the colon indicating the splice point. See *e.g.*, Padgett, *et al.*, 1984. *Ann. Rev. Biochem.* 55:1119-1150. Potential splice sites were identified in order of their matching to this consensus, with a minimum of 4/6 matches beyond the invariant nucleotides.

No ideal 5'-splice site was identified for the GluT1* sequence; thus, the consensus sequence match was relaxed to 3/6 identical nucleotides and one splice site was defined for GluT1*.

Next, potential 3'-intron:exon splice sites were identified based upon the consensus analysis described by Padgett, *et al.* (1984. *Ann. Rev. Biochem.* 55:1119-1150). The sequence between potential 5'-splice sites and the start of the detected interacting sequence was scanned for the invariant AG sequence. The nucleotide preceding was required to be a C or T. Then, nucleotides -5 to -14 from the last intronic G nucleotide had to contain at most two non-T, non-C nucleotides. See *e.g.*, Padgett, *et al.*, 1984. *Ann. Rev. Biochem.* 55:1119-1150.

It should be noted however, that no ideal 3' splice site match was found based on the above analysis for the Set* sequence; accordingly sites containing three non-C, non-T nucleotides upstream of the splice site were then included, resulting in one detected 3' splice site for Set*.

Next, based on the known translational frame of the mature protein and each predicted 5'-splice site, compatible translational frames for successful splicing were defined for potential 3'-splice sites. Nucleic acid sequences can be analyzed by a number of nucleic acid analysis programs available in the art to define possible protein translation products. Translation in the

three forward translation frames was used to define possible open reading frames (contiguous spans of codons for amino acids without the presence of a stop codon). Only 3'-sites that matched the necessary translational frame of a 5'-prime splice junction were retained. Unmatched 5'- or 3'-splice sites were eliminated. For Set*, one possible 5'-splice site and compatible 3'-splice site were located.

Finally, for each possible 5':3' splice site pair, a search for a mammalian branch point consensus sequence was performed. See *e.g.*, Reed and Maniatis, 1988. *Genes Dev.* 2:1268-1276. This sequence (T/C N CTGAC) was selected if matches were found for 5 of the 6 defined nucleotides, and if it was present at the requisite 20-60 nucleotides 5'- to the 3'-splice site. Although not absolutely required for pre mRNA splicing, the efficiency of splicing is related to the presence of the consensus sequence. Thus, 5':3' splice site pairs with a branch point consensus sequence were retained over splice site pairs that did not have a branch point consensus sequence. Branch point consensus sequences were identified for both Set* (Figure 1, callout "D") and GluT1* (Figure 3, callout "D").

Final requirements were that splice variant proteins must encode at least 60 amino acid residues to constitute a viable *in vivo* product. Further, the 3'-terminus of splice variants must, be definition, extend into the identified interacting sequence. The splice variants in this application, Set* and GluT1*, met these requirements.

Splice variant and novel open reading frame sequences were subjected to further searches of the NRDB, a non-redundant compilation of GenBank CDS translations+PDB+ SwissProt+PIR SwissProt sequences, and "month", which includes all new or revised GenBank CDS translation+PDB+ SwissProt+PIR sequences released in the last 30 days, to detect homologies to known protein sequences that were not detected over the span of the known protein sequences. However, no significant homologies to known proteins were detected for CD24*, Set*, GluT1*, or 115392* utilizing this aforementioned analysis.

It should be noted that the present invention is not to be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become apparent to those skilled in the art from the foregoing description and accompanying figures. Such modifications are intended to fall within the scope of the appended claims.

Additionally, various publications are cited herein, the disclosures of which are incorporated by reference in their entireties.

WHAT IS CLAIMED IS:

1. A purified complex of an RB protein and an RB-IP protein, wherein said RB-IP protein is selected from the group consisting of RN-tre, cellular apoptosis susceptibility protein CAS, gamma-interferon inducible protein IP-30, ribosomal protein L6, cytokine IK factor, lactate dehydrogenase B (LDH-B), Nlk1(Nek 2), cyclophilin A, Zap 3, oxidoreductase 17-beta-hydroxysteroid (17 beta-HSD6), CD24*, Set*, glutamate transporter GluT1* or clone 115392*.

2. The purified complex of claim 1, wherein said proteins are human proteins.

3. A purified complex selected from the group consisting of a complex of a derivative of an RB and an RB-IP protein, a complex of a RB protein and a derivative of an RB-IP protein, and a complex of a derivative of an RB protein and a derivative of an RB-IP protein, wherein the derivative of the RB protein possesses the ability to form a complex with a native, wild-type RB-IP protein and the derivative of the RB-IP protein is able to form a complex with a native, wild-type RB protein, in which the RB-IP is selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392*.

4. The purified complex of claim 3, wherein said derivative of the RB protein and/or the RB-IP protein is fluorescently labeled.

5. A chimeric protein comprising a fragment of an RB protein consisting of at least 6 amino acid residues fused, via a covalent bond, to a fragment of an RB-IP protein also consisting of at least 6 amino acid residues.

6. The chimeric protein of claim 5, wherein said fragment of the RB protein is a fragment capable of binding the RB-IP protein and wherein said fragment of the RB-IP protein is a fragment capable of binding the RB protein.

7. The chimeric protein of claim 6, wherein said fragment of the RB protein and said fragment of the RB-IP protein form a RB•RB-IP complex.

8. An antibody which immunospecifically-binds the complex of claim 1, or a fragment or derivative of said antibody containing the binding domain thereof.

9. The antibody of claim 8, which does not immunospecifically-bind an RB protein or an RB-IP protein that is not part of a RB•RB-IP complex.

10. An isolated nucleic acid or an isolated combination of nucleic acids comprising a nucleotide sequence encoding an RB protein and a nucleotide sequence encoding an RB-IP protein selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392*.

11. The isolated nucleic acid or isolated combination of nucleic acids of claim 10 which are nucleic acid vectors.

12. The isolated nucleic acid or isolated combination of nucleic acids of claim 11, wherein the RB protein coding sequence and the RB-IP protein coding sequence are operably-linked to non-native promoters.

13. An isolated nucleic acid that comprises a nucleotide sequence encoding the chimeric protein of claim 7.

14. A recombinant cell containing the nucleic acid of claim 10, wherein said nucleic acid is recombinant.

15. A recombinant cell containing the nucleic acid of claim 12, wherein said nucleic acid is recombinant.

16. A recombinant cell containing the nucleic acid of claim 13, wherein said nucleic acid is recombinant.

17. A purified protein selected from the group consisting of CD24*, Set*, GluT1*, and 115392*.

18. The protein of claim 17, wherein said protein is a human protein.

19. The protein of claim 18 which comprises an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:8, SEQ ID NO:11, and SEQ ID NO:13.

20. A purified protein encoded by a nucleic acid hybridizable to the inverse complement of a DNA having a nucleotide sequence selected from the group consisting of a portion of the nucleotide sequence of SEQ ID NO:33, said portion containing the splice site junction resulting from splicing of the unprocessed *set* mRNA, and a portion of the nucleotide sequence of SEQ ID NO:10, said portion containing the splice site junction resulting from splicing of the unprocessed *GluT1* mRNA.

21. A purified protein encoded by a nucleic acid hybridizable to the inverse complement of a DNA having a nucleotide sequence consisting of a portion of the nucleotide sequence selected from the group consisting of SEQ ID NOS: 7 AND 10.

22. A purified derivative or analog of the protein of claim 17, wherein said derivative or analog possesses the ability to bind RB.

23. The derivative or analog of claim 22 which is able to be bound by an antibody directed against a protein selected from the group consisting of Set* and GluT1* which antibody does not possess the ability to bind either Set or GluT1.

24. The derivative or analog of claim 22 which is able to be bound by an antibody directed against a protein selected from the group consisting of CD24* and 115392*.

25. A purified fragment of the protein of claim 17, wherein said fragment comprises an at least 6 amino acid residue portion of said protein, the amino acid sequence of which portion is not contained in Set or GluT1.

26 A purified protein comprising an amino acid sequence which possesses at least a 60% identity to the protein of claim 17, wherein said percentage of identity is determined over an amino acid sequence of identical size to said protein of claim 17.

27. A chimeric protein comprising a fragment of the protein of claim 17, said fragment consisting of at least 6 amino acid residues and which 6 amino acid residues are not a portion of Set or GluT1, fused via a covalent bond to an amino acid sequence of a second protein, in which the second protein is not said protein of claim 17 and is not Set or GluT1.

28. An antibody which immunospecifically-binds the protein of claim 17 but does not immunospecifically-bind Set or GluT1, or a fragment or derivative of said antibody containing the binding domain thereof.

29. An isolated nucleic acid comprising a nucleotide sequence encoding the protein of claim 17.

30. An isolated nucleic acid comprising the nucleotide sequence of SEQ ID NO:4, SEQ ID NO:7, SEQ ID NO:10 or SEQ ID NO:12.

31. A recombinant cell containing the nucleic acid of claim 30, wherein said nucleic acid is recombinant.

32. A pharmaceutical composition comprising a therapeutically- or prophylactically-effective amount of the complex of claim 1; and a pharmaceutically acceptable carrier.

33. The pharmaceutical composition of claim 32, wherein said proteins are human proteins.

34. A pharmaceutical composition comprising a therapeutically- or prophylactically-effective amount of the complex of claim 3; and a pharmaceutically acceptable carrier.

35. A pharmaceutical composition comprising a therapeutically- or prophylactically-effective amount of the chimeric protein of claim 5; and a pharmaceutically acceptable carrier.

36. A pharmaceutical composition comprising a therapeutically- or prophylactically-effective amount of the chimeric protein of claim 6; and a pharmaceutically acceptable carrier.

37. A pharmaceutical composition comprising a therapeutically- or prophylactically-effective amount of the antibody of claim 8, or a fragment or derivative of said antibody, containing the binding domain thereof; and a pharmaceutically acceptable carrier.

38. A pharmaceutical composition comprising a therapeutically- or prophylactically-effective amount of the antibody of claim 9 or a fragment or derivative of said antibody containing the binding domain thereof; and a pharmaceutically acceptable carrier.

39. A pharmaceutical composition comprising a therapeutically- or prophylactically-effective amount of the nucleic acids or combination of nucleic acids of claim 10; and a pharmaceutically acceptable carrier.

40. A pharmaceutical composition comprising a therapeutically- or prophylactically-effective amount of the nucleic acid of claim 13; and a pharmaceutically acceptable carrier.

41. A pharmaceutical composition comprising a therapeutically- or prophylactically-effective amount of the recombinant cell of claim 15; and a pharmaceutically acceptable carrier.

42. A pharmaceutical composition comprising a therapeutically- or prophylactically-effective amount of the recombinant cell of claim 16; and a pharmaceutically acceptable carrier.

43. A pharmaceutical composition comprising a therapeutically- or prophylactically-effective amount of the protein of claim 17; and a pharmaceutically acceptable carrier.

44. The pharmaceutical composition of claim 42 in which the CD24* protein comprises the amino acid sequence as set forth in SEQ ID NO:5, the Set* protein comprises the amino acid sequence as set forth in SEQ ID NO:8, the GluT1* protein comprises the amino acid sequence as set forth in SEQ ID NO:11, and the 115392* protein comprises the amino acid sequence as set forth in SEQ ID NO:13.

45. A pharmaceutical composition comprising a therapeutically- or prophylactically-effective amount of the antibody of claim 28, or a fragment or derivative of said antibody containing the binding domain thereof; and a pharmaceutically acceptable carrier.

46. A pharmaceutical composition comprising a therapeutically- or prophylactically-effective amount of a nucleic acid comprising a nucleotide sequence encoding the protein of claim 17; and a pharmaceutically acceptable carrier.

47. A pharmaceutical composition comprising a therapeutically- or prophylactically-effective amount of a recombinant cell containing the nucleic acid of claim 29; and a pharmaceutically acceptable carrier.

48. A method of producing a complex of a RB protein and an RB-IP protein comprising growing a recombinant cell containing the nucleic acid of claim 10 such that the encoded RB and RB-IP proteins are expressed and bind to each other, and recovering the expressed complex of the RB protein and the RB-IP protein.

49. A method of producing a protein selected from the group consisting of CD24*, Set*, GluT1*, and 115392* comprising growing a recombinant cell containing a nucleic acid encoding said protein such that the encoded protein is expressed, and recovering the expressed protein.

50. A method of diagnosing or screening for the presence of, or a predisposition for, developing a disease or disorder characterized by an aberrant level of a complex of an RB protein and an RB-IP protein, in which the RB-IP is selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392*, in a subject, comprising measuring the level of said complex, RNA encoding the RB and RB-IP proteins, or functional activity of said complex, in a sample derived from the subject, in which an increase or decrease in the level of said complex, said RNA encoding RB and RB-IP, or functional activity of said complex in the sample, relative to the level of said complex, said RNA encoding RB and RB-IP, or functional activity of said complex, found in an analogous sample not having the disease or disorder or a predisposition for developing the disease or disorder, indicates the presence of the disease or disorder, or a predisposition for developing the disease or disorder.

51. A method of diagnosing or screening for the presence of, or a predisposition for developing, a disease or disorder characterized by an aberrant level of a protein or RNA selected from the group consisting of CD24*, Set*, GluT1* and 115392* protein or RNA in a subject, comprising measuring the level of said protein, said RNA, or the functional activity of said protein, in a sample derived from the subject, in which an increase or decrease in the level of said protein, said RNA, or said functional activity in the sample, relative to the level of said protein,

said RNA, or said functional activity found in an analogous sample not having the disease or disorder, or a predisposition for developing the disease or disorder, indicates the presence of the disease or disorder, or a predisposition for developing the disease or disorder.

52. A kit comprising in one or more containers a substance selected from the group consisting of a complex of an RB and an RB-IP, an antibody against said complex, nucleic acid probes capable of hybridizing to RNA of RB and RNA of said RB-IP, or pairs of nucleic acid primers capable of priming amplification of at least a portion of a gene for RB and a gene for said RB-IP, in which said RB-IP is selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392*.

53. A method of treating or preventing a disease or disorder involving aberrant levels of a complex of RB and an RB-IP, in which the RB-IP is selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392*, in a subject comprising administering to a subject in which such treatment or prevention is desired, a therapeutically effective amount of a molecule or molecules that modulate the function of said complex.

54. The method of claim 53 in which said disease or disorder involves decreased levels of said complex, and said molecule or molecules promote the function of the complex of an RB and an RB-IP and are selected from the group consisting of a complex of RB and an RB-IP; a derivative or analog of a complex of RB and an RB-IP, which complex is more stable or more active than the wild type complex; nucleic acids encoding the RB and an RB-IP protein, and nucleic acids encoding a derivative or analog of RB and RB-IP that form a complex that is more stable or more active than the wild type complex.

55. The method of claim 53 in which said disease or disorder involves increased levels of said complex, and said molecule or molecules inhibit the function of said complex and are selected from the group consisting of an antibody against said complex or a fragment or derivative thereof containing the binding region thereof; *RB* and an *RB-IP* antisense nucleic acids; and nucleic acids comprising at least a portion of an *RB* and an *RB-IP* gene into which a heterologous nucleotide sequence has been inserted such that said heterologous sequence inactivates the biological activity of the RB and RB-IP genes, in which the *RB* and the *RB-IP*

gene portions flank the heterologous sequences so as to promote homologous recombination with genomic *RB* and *RB-IP* genes.

56. A method of treating or preventing a disease or disorder involving an aberrant level of an RB-IP selected from the group consisting of CD24*, Set*, GluT1* and 115392* in a subject, comprising administering to a subject in which such treatment or prevention is desired a therapeutically effective amount of a molecule that modulates the function of said RB-IP.

57. The method of claim 56 in which said disease or disorder involves a decreased level of the RB-IP and said molecule promotes the function of the RB-IP and is selected from the group consisting of the RB-IP protein, a derivative or analog of the RB-IP that is active in binding RB, a nucleic acid encoding the RB-IP protein, or a nucleic acid encoding a derivative or analog of the RB-IP that is active in binding RB.

58. The method of claim 56 in which said disease or disorder involves an increased level of the RB-IP, and said molecule inhibits the RB-IP function and is selected from the group consisting of an anti-RB-IP antibody or a fragment or derivative thereof containing the binding region thereof, an *RB-IP* antisense nucleic acid, and a nucleic acid comprising at least a portion of the *RB-IP* gene into which a heterologous nucleotide sequence has been inserted such that said heterologous sequence inactivates the biological activity of the *RB-IP* gene, in which the *RB-IP* gene portion flanks the heterologous sequence so as to promote homologous recombination with a genomic *RB-IP* gene.

59. A method of screening a purified complex of RB and a RB-IP selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392*, or a derivative of said complex, or a modulator of the activity of said complex, for anti-cancer activity comprising measuring the survival or proliferation of cells from a cell line which is derived from or displays characteristics associated with malignant disorder, which cells have been contacted with the complex, derivative, or modulator; and comparing the survival or proliferation in the cells contacted with the complex, derivative or modulator with said survival or proliferation in cells not so contacted, wherein a lower level in said contacted cells indicates that the complex, derivative or modulator has anti-tumor activity.

60. A method of screening a purified complex of RB and a RB-IP selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392* or a derivative of said complex, or a modulator of the activity of said complex for anti-cancer activity by a method comprising administering the complex, derivative or modulator to a test animal, which test animal has a tumor, or has increased propensity to develop a tumor, or which test animal does not have a tumor and is subsequently challenged with tumor cells or tumorigenic agents; and measuring tumor growth or regression in said test animal, wherein decreased tumor growth or increased tumor regression, failure to develop tumors, or prevention of tumor growth in test animals administered said complex, derivative or modulator compared to test animals not so administered indicates that the complex, derivative or modulator has anti-cancer activity.

61. A method for screening a purified complex of RB and an RB-IP selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392* or a derivative of said complex, or a modulator of the activity of said complex for activity in treating or preventing neurodegenerative disease comprising contacting cultured cells that exhibit an indicator of a neurodegenerative disease, *in vitro* with said complex, derivative or modulator; and comparing the level of said indicator in the cells contacted with the complex, derivative, or modulator with said level of said indicator in cells not so contacted, wherein a lower level in said contacted cells indicates that the complex, derivative or modulator has activity in treating or preventing neurodegenerative disease.

62. A method for screening a purified complex of RB and an RB-IP selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392*, or a derivative of said complex, or a modulator of the activity of said complex, for activity in treating or preventing neurodegenerative disease comprising administering said complex, derivative or modulator to a test animal, which test animal exhibits symptoms of a neurodegenerative disease or which test animal is predisposed to develop symptoms of a neurodegenerative disease; and measuring the change in said symptoms of the neurodegenerative disease after administration of said complex, derivative, or modulator, wherein a reduction in the severity of the symptoms of the neurodegenerative disease or prevention of the symptoms of the neurodegenerative disease indicates that the complex, derivative or modulator has activity in treating or preventing neurodegenerative disease.

63. A method for screening a purified complex of RB and an RB-IP selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392* or a derivative of said complex, or a modulator of the activity of said complex for activity in treating or preventing viral disease comprising contacting cultured cells that exhibit an indicator of a neurodegenerative disease, *in vitro* with said complex, derivative or modulator; and comparing the level of said indicator in the cells contacted with the complex, derivative, or modulator with said level of said indicator in cells not so contacted, wherein a lower level in said contacted cells indicates that the complex, derivative or modulator has activity in treating or preventing viral disease.

64. A method for screening a purified complex of RB and an RB-IP selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392*, or a derivative of said complex, or a modulator of the activity of said complex, for activity in treating or preventing viral disease comprising administering said complex, derivative or modulator to a test animal, which test animal exhibits symptoms of a viral disease or which test animal is predisposed to develop symptoms of a viral disease; and measuring the change in said symptoms of viral disease after administration of said complex, derivative, or modulator, wherein a reduction in the severity of the symptoms of the viral disease or prevention of the symptoms of the viral disease indicates that the complex, derivative or modulator has activity in treating or preventing viral disease.

65. A method of screening for a molecule that modulates directly or indirectly the formation of a complex of RB and a RB-IP, in which said RB-IP is selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392*, comprising measuring the levels of said complex formed from RB and RB-IP proteins in the presence of said molecule under conditions conducive to formation of the complex; and comparing the levels of said complex with the levels of said complex that are formed in the absence of said molecule, wherein a lower or higher level of said complex in the presence of said molecule indicates that the molecule modulates formation of said complex.

66. A recombinant, non-human animal in which both an endogenous *RB* gene and an endogenous *RB-IP* gene selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392* have been deleted or

inactivated by homologous recombination or insertional mutagenesis of said animal or an ancestor thereof.

67. A recombinant, non-human animal containing both an *RB* gene and an *RB-IP* gene selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392* in which the *RB* gene is under the control of a promoter that is not the native *RB* gene promoter and the *RB-IP* gene is under the control of a promoter that is not the native *RB-IP* gene promoter.

68. A recombinant, non-human animal containing a transgene comprising a nucleic acid sequence encoding the chimeric protein of claim 6.

69. A method of modulating the activity or levels of RB by contacting a cell with, or administering an animal expressing a *RB* gene, a protein selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392*, or a nucleic acid encoding said protein, or an antibody that immunospecifically-binds said protein, or a fragment or derivative of said antibody containing the binding domain thereof.

70. A method of modulating the activity or levels of a protein selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392* by contacting a cell with, or administering an animal expressing a gene encoding said protein, RB, or a nucleic acid encoding RB, or an antibody that immunospecifically-binds RB, or a fragment or derivative of said antibody containing the binding domain thereof.

71. A method of modulating the activity or levels of a complex of RB and a protein selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392*, by contacting a cell with, or administering an animal expressing and forming said complex, a molecule that modulates the formation of said complex.

72. A method for identifying a molecule that modulates activity of RB or a protein selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2),

cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392* or a complex of RB and said protein comprising contacting one or more candidate molecules with RB in the presence of said protein; and measuring the amount of complex that forms between RB and said protein; wherein an increase or decrease in the amount of complex that forms relative to the amount that forms in the absence of the candidate molecules indicates that the molecules modulate the activity of RB or said protein or said complex of RB and said protein.

73. The method of claim 72, wherein said contacting is carried out by administering the candidate molecules to the recombinant, non-human animal of claim 67.

74. The method of claim 72, wherein said contacting is carried out *in vitro*; and RB, said protein, and said candidate molecules are purified.

75. A method for screening a derivative or analog of RB for biological activity comprising contacting said derivative or analog of RB with a protein selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392*, and detecting the formation of a complex between said derivative or analog of RB and said protein; wherein the formation of said complex indicates that said derivative or analog of RB has biological activity.

76. A method for screening a derivative or analog of a protein selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392*, for biological activity comprising contacting said derivative or analog of said protein with RB; and detecting the formation of a complex between said derivative or analog of said protein and RB; wherein the formation of said complex indicates that said derivative or analog of said protein has biological activity.

77. A method of monitoring the efficacy of a treatment of a disease or disorder characterized by an aberrant level of a complex of RB protein and a RB-IP protein, in a subject administered said treatment, for said disease or disorder comprising measuring the level of said complex, RNA encoding the RB and RB-IP proteins, or functional activity of said complex, in a sample derived from said subject wherein said sample is taken from said subject after the administration of said treatment, and compared to: (*i*) said level in a sample taken from said subject prior to the administration of the treatment or (*ii*) a standard level associated with the

pretreatment stage of the disease or disorder, in which the change, or lack of change in the level of said complex, said RNA encoding RB and RB-IP, or functional activity of said complex, in said sample taken after the administration of said treatment relative to the level of said complex, said RNA encoding RB and RB-IP or functional activity of said complex in said sample taken before the administration of said treatment or to said standard level, indicates whether said administration is effective for treating said disease or disorder.

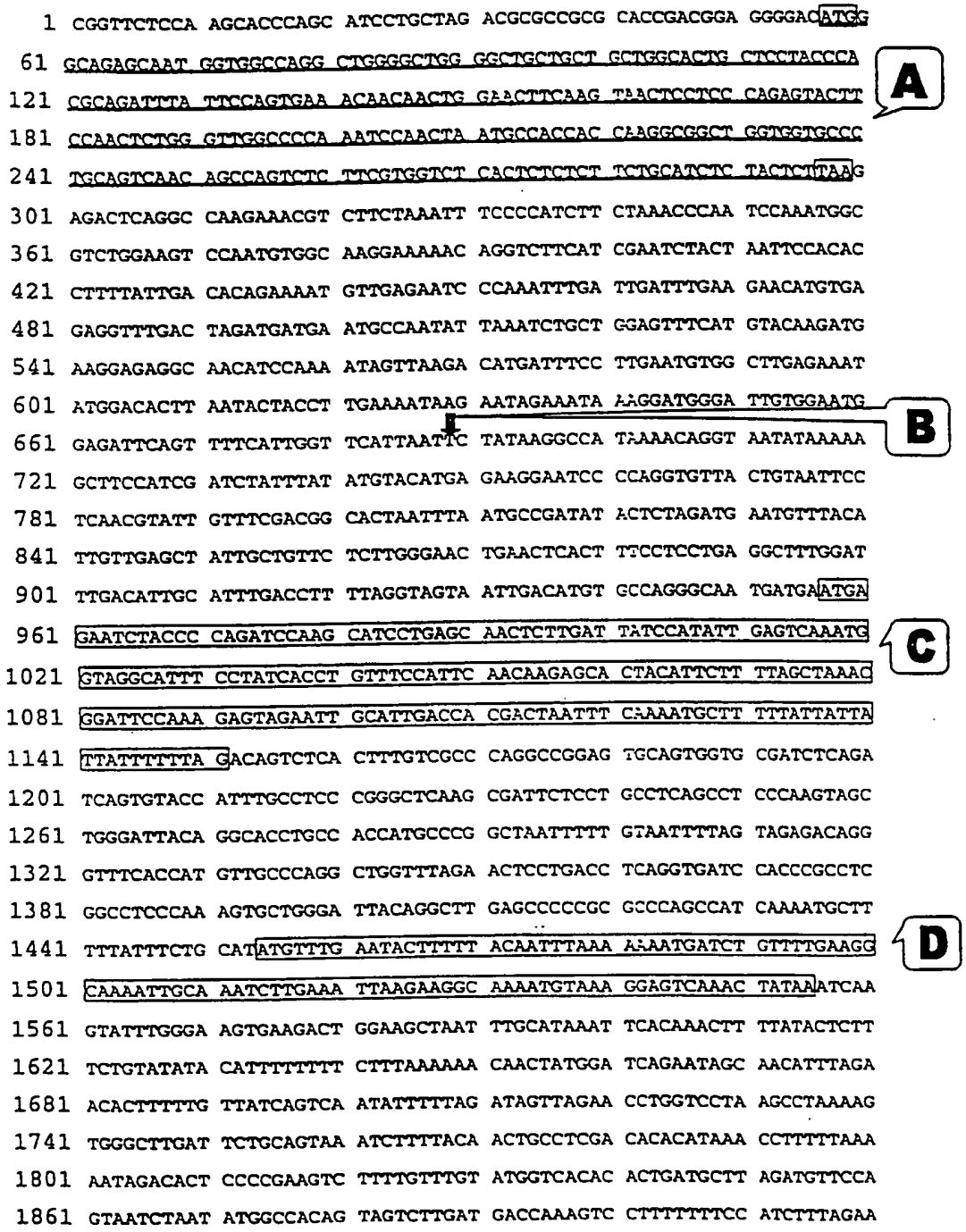
78. A method of treating or preventing cancer or a cell proliferation disorder in a subject comprising administering to a subject in which such treatment or prevention is desired a therapeutically effective amount of a molecule that modulates the function of a complex of RB and a RB-IP protein selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392*, or a combination of any one or more of the foregoing.

79. A method of treating or preventing neurodegenerative disease in a subject comprising administering to a subject in which such treatment or prevention is desired a therapeutically effective amount of a molecule that modulates the function of a complex of RB and an RB-IP protein selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392*, or a combination of any one or more of the foregoing.

80. A method of treating or preventing viral disease in a subject comprising administering to a subject in which such treatment or prevention is desired a therapeutically effective amount of a molecule that modulates the function of a complex of RB and an RB-IP protein selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392*, or a combination of any one or more of the foregoing.

81. A purified fragment of a protein selected from the group consisting of RN-tre, CAS, IP-30, L6, IK, LDH-B, Nlk1(Nek2), cyclophilin A, Zap 3, CD24*, Set*, GluT1*, and 115392*, which fragment binds RB.

Figure 1




```
   1 CGGTTCTCCA AGCACCCAGC ATCCTGCTAG ACGCGCCGCG CACCGACGGA GGGGACATGG
  61 GCAGAGCAAT GGTGGCCAGG CTGGGGCTGG GGCTGCTGCT GCTGGCACTG CTCCTACCCA
 121 CGCAGATTTA TTCCAGTGAA ACAACAACTG GAACTTCAAG TAACTCCTCC CAGAGTACTT
 181 CCAACTCTGG GTTGGCCCCA AATCCAACTA ATGCCACCAC CAAGGCGGCT GGTGGTGCCC
 241 TGCAGTCAAC AGCCAGTCTC TTCGTGGTCT CACTCTCTCT TCTGCATCTC TACTCTTAAG
 301 AGACTCAGGC CAAGAAACGT CTTCTAAATT TCCCCATCTT CTAAACCCAA TCCAAATGGC
 361 GTCTGGAAGT CCAATGTGGC AAGGAAAAAC AGGTCTTCAT CGAATCTACT AATTCCACAC
 421 CTTTTATTGA CACAGAAAAT GTTGAGAATC CCAAATTTGA TTGATTTGAA GAACATGTGA
 481 GAGGTTTGAC TAGATGATGA ATGCCAATAT TAAATCTGCT GGAGTTTCAT GTACAAGATG
 541 AAGGAGAGGC AACATCCAAA ATAGTTAAGA CATGATTTCC TTGAATGTGG CTTGAGAAAT
 601 ATGGACACTT AATACTACCT TGAAAATAAG AATAGAAATA AAGGATGGGA TTGTGGAATG
 661 GAGATTCAGT TTTCATTGGT TCATTAATTC TATAAGGCCA TAAAACAGGT AATATAAAAA
 721 GCTTCCATCG ATCTATTTAT ATGTACATGA GAAGGAATCC CCAGGTGTTA CTGTAATTCC
 781 TCAACGTATT GTTTCGACGG CACTAATTTA ATGCCGATAT ACTCTAGATG AATGTTTACA
 841 TTGTTGAGCT ATTGCTGTTC TCTTGGGAAC TGAACTCACT TTCCTCCTGA GGCTTTGGAT
 901 TTGACATTGC ATTTGACCTT TTAGGTAGTA ATTGACATGT GCCAGGGCAA TGATGAATGA
 961 GAATCTACCC CAGATCCAAG CATCCTGAGC AACTCTTGAT TATCCATATT GAGTCAAATG
1021 GTAGGCATTT CCTATCACCT GTTTCCATTC AACAAGAGCA CTACATTCTT TTAGCTAAAC
1081 GGATTCCAAA GAGTAGAATT GCATTGACCA CGACTAATTT CAAAATGCTT TTTATTATTA
1141 TTATTTTTTA GACAGTCTCA CTTTGTCGCC CAGGCCGGAG TGCAGTGGTG CGATCTCAGA
1201 TCAGTGTACC ATTTGCCTCC CGGGCTCAAG CGATTCTCCT GCCTCAGCCT CCCAAGTAGC
1261 TGGGATTACA GGCACCTGCC ACCATGCCCG GCTAATTTTT GTAATTTTAG TAGAGACAGG
1321 GTTTCACCAT GTTGCCCAGG CTGGTTTAGA ACTCCTGACC TCAGGTGATC CACCCGCCTC
1381 GGCCTCCCAA AGTGCTGGGA TTACAGGCTT GAGCCCCCGC GCCCAGCCAT CAAAATGCTT
1441 TTTATTTCTG CATATGTTTG AATACTTTTT ACAATTTAAA AAAATGATCT GTTTTGAAGG
1501 CAAAATTGCA AATCTTGAAA TTAAGAAGGC AAAATGTAAA GGAGTCAAAC TATAAATCAA
1561 GTATTTGGGA AGTGAAGACT GGAAGCTAAT TTGCATAAAT TCACAAACTT TTATACTCTT
1621 TCTGTATATA CATTTTTTTT CTTTAAAAAA CAACTATGGA TCAGAATAGC AACATTTAGA
1681 ACACTTTTTG TTATCAGTCA ATATTTTTAG ATAGTTAGAA CCTGGTCCTA AGCCTAAAAG
1741 TGGGCTTGAT TCTGCAGTAA ATCTTTTACA ACTGCCTCGA CACACATAAA CCTTTTTAAA
1801 AATAGACACT CCCCGAAGTC TTTTGTTTGT ATGGTCACAC ACTGATGCTT AGATGTTCCA
1861 GTAATCTAAT ATGGCCACAG TAGTCTTGAT GACCAAAGTC CTTTTTTTCC ATCTTTAGAA
```

Figure 1

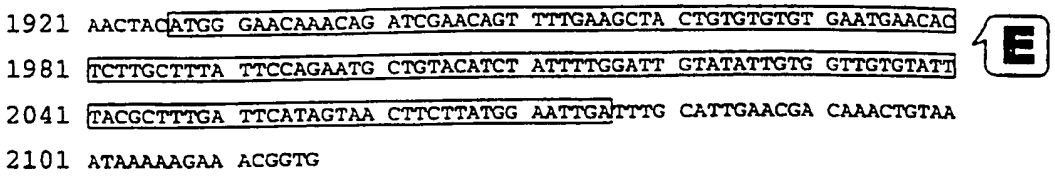

1921 AACTACATGG GAACAAACAG ATCGAACAGT TTTGAAGCTA CTGTGTGTGT GAATGAACAC
1981 TCTTGCTTTA TTCCAGAATG CTGTACATCT ATTTTGGATT GTATATTGTG GTTGTGTATT
2041 TACGCTTTGA TTCATAGTAA CTTCTTATGG AATTGATTTG CATTGAACGA CAAACTGTAA
2101 ATAAAAAGAA ACGGTG
E

Figure 2

A.

```
 57 ATG GGC AGA GCA ATG GTG GCC AGG CTG GGG CTG GGG CTG CTG CTG
    Met Gly Arg Ala Met Val Ala Arg Leu Gly Leu Gly Leu Leu Leu
     1               5                  10                  15

102 CTG GCA CTG CTC CTA CCC ACG CAG ATT TAT TCC AGT GAA ACA ACA
    Leu Ala Leu Leu Leu Pro Thr Gln Ile Tyr Ser Ser Glu Thr Thr
                    20                  25                  30

147 ACT GGA ACT TCA AGT AAC TCC TCC CAG AGT ACT TCC AAC TCT GGG
    Thr Gly Thr Ser Ser Asn Ser Ser Gln Ser Thr Ser Asn Ser Gly
                    35                  40                  45

192 TTG GCC CCA AAT CCA ACT AAT GCC ACC ACC AAG GCG GCT GGT GGT
    Leu Ala Pro Asn Pro Thr Asn Ala Thr Thr Lys Ala Ala Gly Gly
                    50                  55                  60

237 GCC CTG CAG TCA ACA GCC AGT CTC TTC GTG GTC TCA CTC TCT CTT
    Ala Leu Gln Ser Thr Ala Ser Leu Phe Val Val Ser Leu Ser Leu
                    65                  70                  75

282 CTG CAT CTC TAC TCT TAA  299
    Leu His Leu Tyr Ser *
                    80
```

B.

```
957 ATG AGA ATC TAC CCC AGA TCC AAG CAT CCT GAG CAA CTC TTG ATT
    Met Arg Ile Tyr Pro Arg Ser Lys His Pro Glu Gln Leu Leu Ile
     1               5                  10                  15

1002 ATC CAT ATT GAG TCA AAT GGT AGG CAT TTC CTA TCA CCT GTT TCC
    Ile His Ile Glu Ser Asn Gly Arg His Phe Leu Ser Pro Val Ser
                    20                  25                  30
```

Figure 2

```
1047 ATT CAA CAA GAG CAC TAC ATT CTT TTA GCT AAA CGG ATT CCA AAG
     Ile Gln Gln Glu His Tyr Ile Leu Leu Ala Lys Arg Ile Pro Lys
                     35                  40                  45

1092 AGT AGA ATT GCA TTG ACC ACG ACT AAT TTC AAA ATG CTT TTT ATT
     Ser Arg Ile Ala Leu Thr Thr Thr Asn Phe Lys Met Leu Phe Ile
                     50                  55                  60

1137 ATT ATT ATT TTT TAG  1151
     Ile Ile Ile Phe *
                 64
```

Figure 3

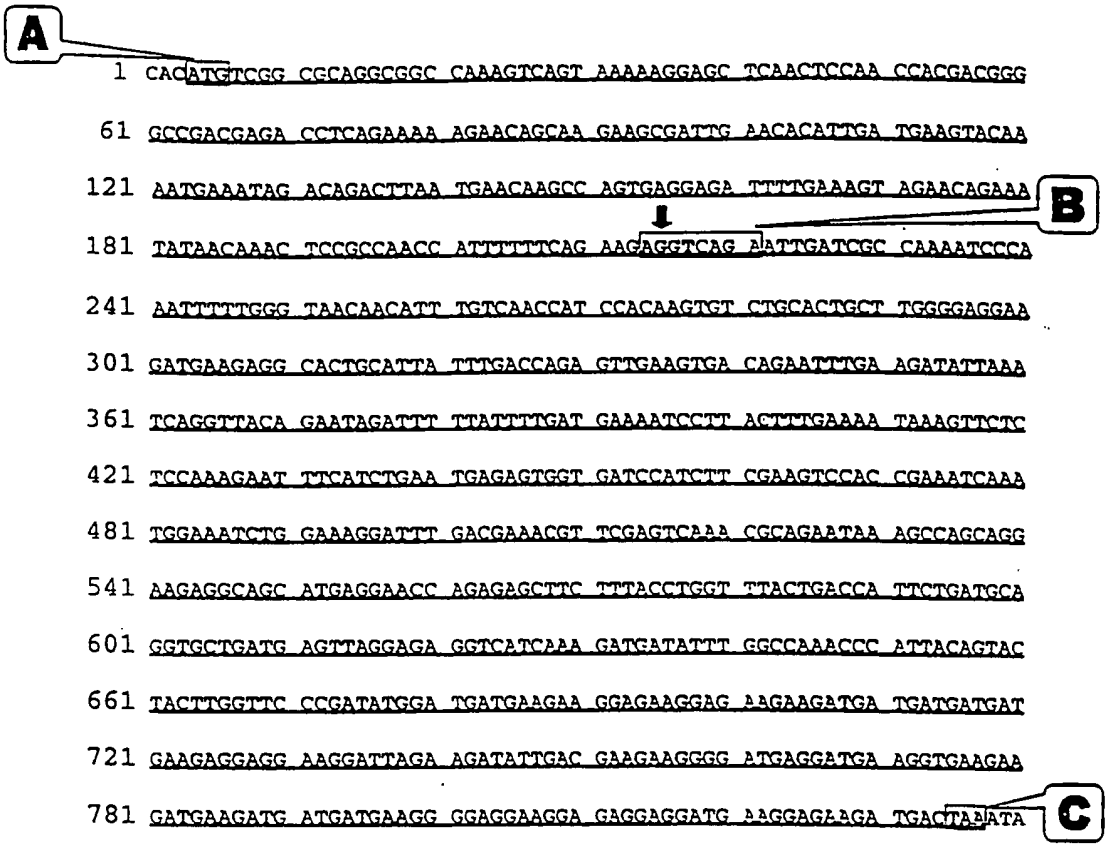

A
B
C
1 CACATGTCGG CGCAGGCGGC CAAAGTCAGT AAAAAGGAGC TCAACTCCAA CCACGACGGG
61 GCCGACGAGA CCTCAGAAAA AGAACAGCAA GAAGCGATTG AACACATTGA TGAAGTACAA
121 AATGAAATAG ACAGACTTAA TGAACAAGCC AGTGAGGAGA TTTTGAAAGT AGAACAGAAA
181 TATAACAAAC TCCGCCAACC ATTTTTTCAG AAGAGGTCAG AATTGATCGC CAAAATCCCA
241 AATTTTTGGG TAACAACATT TGTCAACCAT CCACAAGTGT CTGCACTGCT TGGGGAGGAA
301 GATGAAGAGG CACTGCATTA TTTGACCAGA GTTGAAGTGA CAGAATTTGA AGATATTAAA
361 TCAGGTTACA GAATAGATTT TTATTTTGAT GAAAATCCTT ACTTTGAAAA TAAAGTTCTC
421 TCCAAAGAAT TTCATCTGAA TGAGAGTGGT GATCCATCTT CGAAGTCCAC CGAAATCAAA
481 TGGAAATCTG GAAAGGATTT GACGAAACGT TCGAGTCAAA CGCAGAATAA AGCCAGCAGG
541 AAGAGGCAGC ATGAGGAACC AGAGAGCTTC TTTACCTGGT TTACTGACCA TTCTGATGCA
601 GGTGCTGATG AGTTAGGAGA GGTCATCAAA GATGATATTT GGCCAAACCC ATTACAGTAC
661 TACTTGGTTC CCGATATGGA TGATGAAGAA GGAGAAGGAG AAGAAGATGA TGATGATGAT
721 GAAGAGGAGG AAGGATTAGA AGATATTGAC GAAGAAGGGG ATGAGGATGA AGGTGAAGAA
781 GATGAAGATG ATGATGAAGG GGAGGAAGGA GAGGAGGATG AAGGAGAAGA TGACTAAATA
841 GAACACTGAT GGATTCCAAC CTTCCTTTTT TTAAATTTTC TCCAGTCCCT GGGAGCAAGT
901 TGCAGTCTTT TTTTTTTTTT TTTTTTTTTT CCCTCTTGTG CTCAGTCGCC CTGTTCTTGA
961 GGTCTCTTTT CTCTACTCCA TGGTTCTCAA TTTATTTGGG GGGAAATACC TTGAGCAGAA
1021 TACAATGGGA AAAGAGTCTC TACCCCTTTC TGTTCGAAGT TCATTTTTAT CCCTTCCTGT
1081 CTGAACAAAA ACTGTATGGA ATCAACACCA CCGAGCTCTG TGGGAAAAAA GAAAAACCTG
1141 CTCCCTTTGC TCTGCTGGAA GCTGGAGGGT GCTAGGCCCC TGTGTAGTAG TGTATAGAAT
1201 TCTAGCTTTT TTCCTCCTTT CTCTGTATAT TGGGCTCAGA GAGTACACTG TGTCTCTATG
1261 TGAATATGGA CAGTTAGCAT TTACCAACAT GTATCTGTCT ACTTTCTCTT GTTTAAAAAA
1321 AGAAAAAAAA ACTTAAAAAA ATGGGGTTAT AGAAGGTCAG CAAAGGGGTG GGGTTTGAGA
1381 TGTTTGGGTG GGTTAGTGGG CATTTTGACA ACATGGCTTC TCCTTTGGCA TGTTTAATTG
1441 TGATATTTGA CAGACATCCT TGCAGTTTAA GATGACACTT TTAAAATAAA TTCTCTCCTA

Figure 3

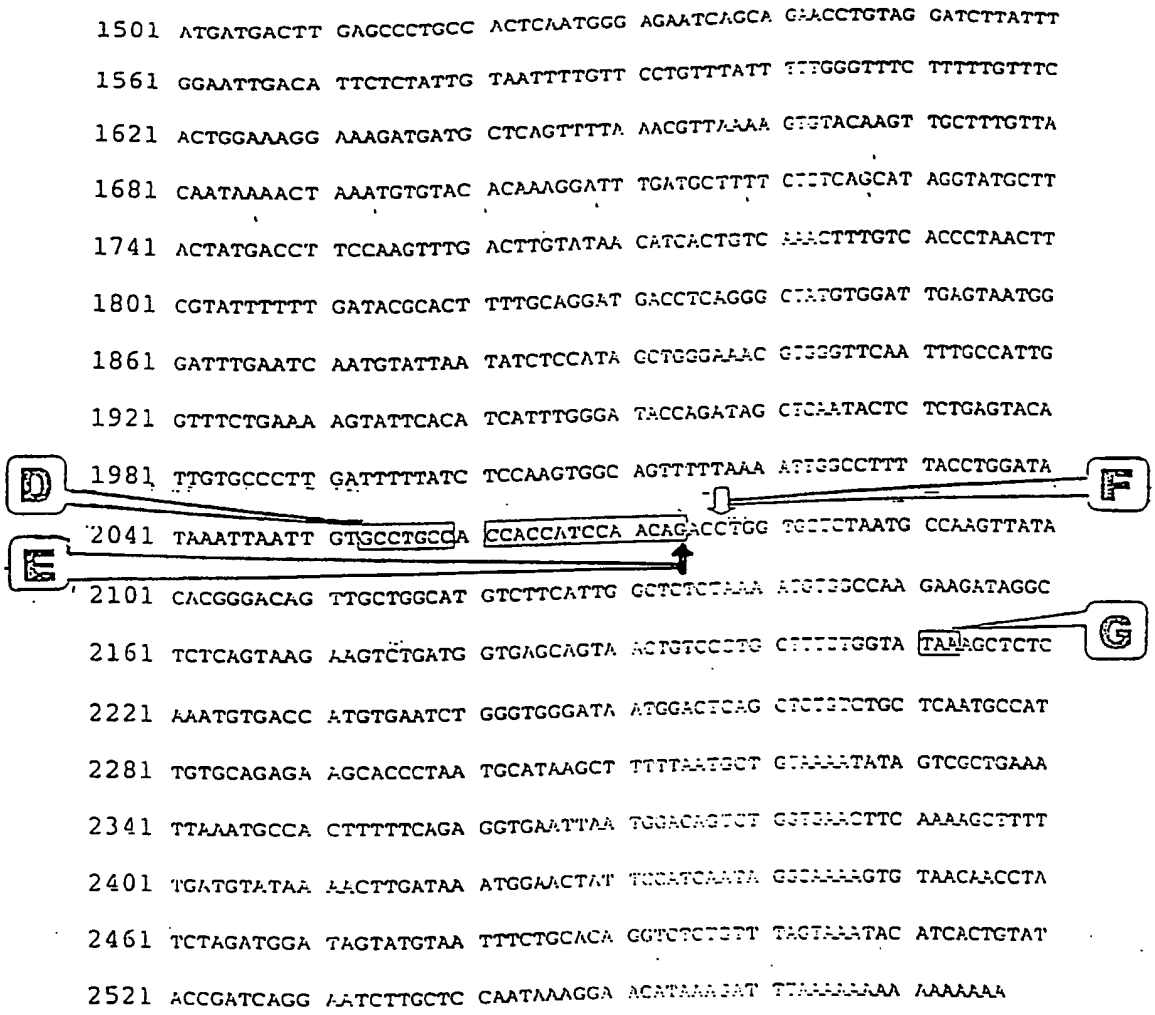

1501 ATGATGACTT GAGCCCTGCC ACTCAATGGG AGAATCAGCA GAACCTGTAG GATCTTATTT
1561 GGAATTGACA TTCTCTATTG TAATTTTGTT CCTGTTTATT TTTGGGTTTC TTTTTGTTTC
1621 ACTGGAAAGG AAAGATGATG CTCAGTTTTA AACGTTAAAA GTGTACAAGT TGCTTTGTTA
1681 CAATAAAACT AAATGTGTAC ACAAAGGATT TGATGCTTTT CTTTCAGCAT AGGTATGCTT
1741 ACTATGACCT TCCAAGTTTG ACTTGTATAA CATCACTGTC AAACTTTGTC ACCCTAACTT
1801 CGTATTTTTT GATACGCACT TTTGCAGGAT GACCTCAGGG CTATGTGGAT TGAGTAATGG
1861 GATTTGAATC AATGTATTAA TATCTCCATA GCTGGGAAAC GTGGGTTCAA TTTGCCATTG
1921 GTTTCTGAAA AGTATTCACA TCATTTGGGA TACCAGATAG CTCAATACTC TCTGAGTACA
1981 TTGTGCCCTT GATTTTTATC TCCAAGTGGC AGTTTTTAAA ATTGGCCTTT TACCTGGATA
2041 TAAATTAATT GTGCCTGCCA CCACCATCCA ACAGACCTGG TGCTCTAATG CCAAGTTATA
2101 CACGGGACAG TTGCTGGCAT GTCTTCATTG GCTCTCTAAA ATGTGGCCAA GAAGATAGGC
2161 TCTCAGTAAG AAGTCTGATG GTGAGCAGTA ACTGTCCCTG CTTTTTGGTA TAAAGCTCTC
2221 AAATGTGACC ATGTGAATCT GGGTGGGATA ATGGACTCAG CTCTGTCTGC TCAATGCCAT
2281 TGTGCAGAGA AGCACCCTAA TGCATAAGCT TTTTAATGCT GTAAAATATA GTCGCTGAAA
2341 TTAAAATGCCA CTTTTTCAGA GGTGAATTAA TGGACAGTCT GGTGAACTTC AAAAGCTTTT
2401 TGATGTATAA AACTTGATAA ATGGAACTAT TGCATCAATA GGCAAAAGTG TAACAACCTA
2461 TCTAGATGGA TAGTATGTAA TTTCTGCACA GGTCTCTGTT TAGTAAATAC ATCACTGTAT
2521 ACCGATCAGG AATCTTGCTC CAATAAAGGA ACATAAAAAT TTAAAAAAAA AAAAAAA
D
E
F
G

Figure 4

```
 1 ATG TCG GCG CAG GCG GCC AAA GTC AGT AAA AAG GAG CTC AAC TCC
   Met Ser Ala Gln Ala Ala Lys Val Ser Lys Lys Glu Leu Asn Ser
    1               5                  10                  15

46 AAC CAC GAC GGG GCC GAC GAG ACC TCA GAA AAA GAA CAG CAA GAA
   Asn His Asp Gly Ala Asp Glu Thr Ser Glu Lys Glu Gln Gln Glu
                   20                  25                  30

91 GCG ATT GAA CAC ATT GAT GAA GTA CAA AAT GAA ATA GAC AGA CTT
   Ala Ule Glu His Ile Asp Glu Val Gln Asn Glu Ile Asp Arg Leu
                   35                  40                  45

136 AAT GAA CAA GCC AGT GAG GAG ATT TTG AAA GTA GAA CAG AAA TAT
    Asn Glu Gln Als Ser Glu Glu Ile Leu Lys Val Glu Gln Lys Tyr
                    50                  55                  60
                                        (A)  ↓   ↓ (B)
181 AAC AAA CTC CGC CAA CCA TTT TTT CAG AAG AGA CCT GGT GCT CTA
    Asn Lys Leu Arg Gln Pro Phe Phe Gln Lys Arg Pro Gly Ala Leu
                    65                  70                  75

226 ATG CCA AGT TAT ACA CGG GAC AGT TGC TGG CAT GTC TTC ATT GGC
    Met Pro Ser Tyr Thr Arg Asp Ser Cys Trp His Val Phe Ile Gly
                    80                  85                  90

271 TCT CTA AAA TGT GGC CAA GAA GAT AGG CTC TCA GTA AGA AGT CTG
    Ser Leu Lys Cys Gly Gln Glu Asp Arg Leu Ser Val Arg Ser Leu
                    95                  100                 105

316 ATG GTG AGC AGT AAC TGT CCC TGC TTT CTG GTA TAA  351
    Met Val Ser Ser Asn Cyc Pro Cys Phe Leu Val *
                    110                 115 116
```

Figure 5

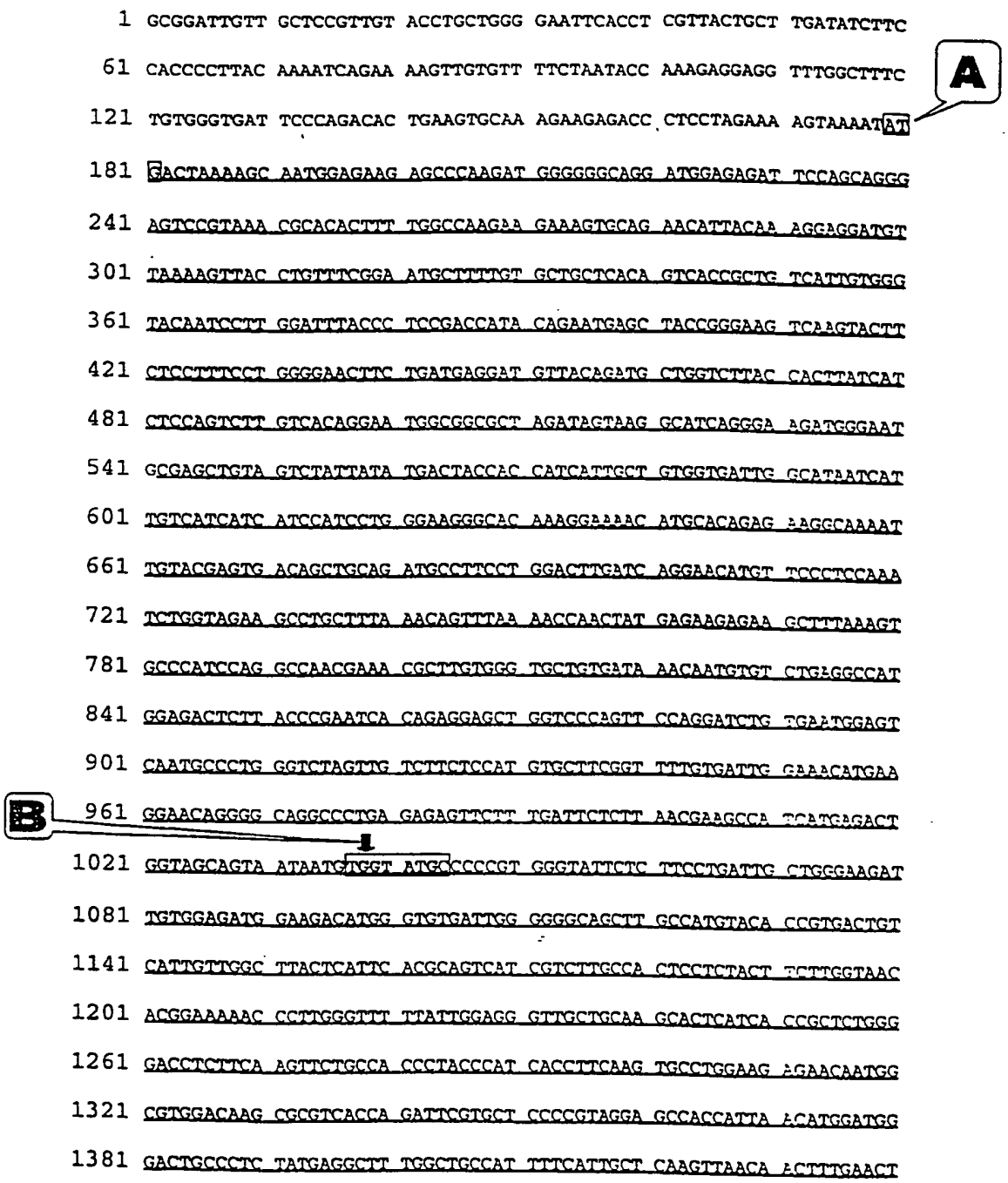

A
B
1 GCGGATTGTT GCTCCGTTGT ACCTGCTGGG GAATTCACCT CGTTACTGCT TGATATCTTC
61 CACCCCTTAC AAAATCAGAA AAGTTGTGTT TCTAATACC AAAGAGGAGG TTTGGCTTTC
121 TGTGGGTGAT TCCCAGACAC TGAAGTGCAA AGAAGAGACC CTCCTAGAAA AGTAAAATAT
181 GACTAAAAGC AATGGAGAAG AGCCCAAGAT GGGGGGCAGG ATGGAGAGAT TCCAGCAGGG
241 AGTCCGTAAA CGCACACTTT TGGCCAAGAA GAAAGTGCAG AACATTACAA AGGAGGATGT
301 TAAAAGTTAC CTGTTTCGGA ATGCTTTTGT GCTGCTCACA GTCACCGCTG TCATTGTGGG
361 TACAATCCTT GGATTTACCC TCCGACCATA CAGAATGAGC TACCGGGAAG TCAAGTACTT
421 CTCCTTTCCT GGGGAACTTC TGATGAGGAT GTTACAGATG CTGGTCTTAC CACTTATCAT
481 CTCCAGTCTT GTCACAGGAA TGGCGGCGCT AGATAGTAAG GCATCAGGGA AGATGGGAAT
541 GCGAGCTGTA GTCTATTATA TGACTACCAC CATCATTGCT GTGGTGATTG GCATAATCAT
601 TGTCATCATC ATCCATCCTG GGAAGGGCAC AAAGGAAAAC ATGCACAGAG AAGGCAAAAT
661 TGTACGAGTG ACAGCTGCAG ATGCCTTCCT GGACTTGATC AGGAACATGT TCCCTCCAAA
721 TCTGGTAGAA GCCTGCTTTA AACAGTTTAA AACCAACTAT GAGAAGAGAA GCTTTAAAGT
781 GCCCATCCAG GCCAACGAAA CGCTTGTGGG TGCTGTGATA AACAATGTGT CTGAGGCCAT
841 GGAGACTCTT ACCCGAATCA CAGAGGAGCT GGTCCCAGTT CCAGGATCTG TGAATGGAGT
901 CAATGCCCTG GGTCTAGTTG TCTTCTCCAT GTGCTTCGGT TTTGTGATTG GAAACATGAA
961 GGAACAGGGG CAGGCCCTGA GAGAGTTCTT TGATTCTCTT AACGAAGCCA TCATGAGACT
1021 GGTAGCAGTA ATAATGTGGT ATGCCCCCGT GGGTATTCTC TTCCTGATTG CTGGGAAGAT
1081 TGTGGAGATG GAAGACATGG GTGTGATTGG GGGGCAGCTT GCCATGTACA CCGTGACTGT
1141 CATTGTTGGC TTACTCATTC ACGCAGTCAT CGTCTTGCCA CTCCTCTACT TCTTGGTAAC
1201 ACGGAAAAAC CCTTGGGTTT TTATTGGAGG GTTGCTGCAA GCACTCATCA CCGCTCTGGG
1261 GACCTCTTCA AGTTCTGCCA CCCTACCCAT CACCTTCAAG TGCCTGGAAG AGAACAATGG
1321 CGTGGACAAG CGCGTCACCA GATTCGTGCT CCCCGTAGGA GCCACCATTA ACATGGATGG
1381 GACTGCCCTC TATGAGGCTT TGGCTGCCAT TTTCATTGCT CAAGTTAACA ACTTTGAACT

Figure 5

1441 GAACTTCGGA CAAATTATTA CAATCAGCAT CACAGCCACA GCTGCCAGTA TTGGGGCAGC

1501 TGGAATTCCT CAGGCGGGCC TGGTCACTAT GGTCATTGTG CTGACATCTG TCGGCCTGCC

1561 CACTGACGAC ATCACGCTCA TCATCGCGGT GGACTGGTTC CTGGATCGCC TCCGGACCAC

1621 CACCAACGTA CTGGGAGACT CCCTGGGAGC TGGGATTGTG GAGCACTTGT CACGACATGA

1681 ACTGAAGAAC AGAGATGTTG AAATGGGTAA CTCAGTGATT GAAGAGAATG AAATGAAGAA

1741 ACCATATCAA CTGATTGCAC AGGACAATGA AACTGAGAAA CCCATCGACA GTGAAACCAA

1801 GATGTAGACT AACATAAAGA AACACTTTCT TGAGCACCAG GTGTTAAAAA CCATTATAAA

1861 ATCTTTCCAT CTCATTACAG CTCATTCGCT CCAGCAAGCC CGTCATCTTC CCTTTCCTCC

1921 CTTCTGATAA GACTGGAAAA TAGTCCTCCA AAACACAAGG GAGGATTTTG GGTGGCCAAA

1981 GTGTACAATT TTCATCCCAC AATTGAAATT TTTAAATCAT TTCATGTTAG TCTTACCGAA

2041 TAAGGTACCA AGATCACAAA TAGTGTTGAT CAGATCTTAC AAGTTTATGT GGCACACAAT

2101 TCCTATAAAT GTGATTTTTT TATATAAGTT AAAGAGACAA ATAGTAGGCT AAAAACATTT

2161 TAAAATCAAC TTTTGAAATT TAAAAATCTT TCAGAATACA ATTCAGTTTT AGTTTCAAAA

2221 TGTTAACAAC TTGAATTACA ACCGGTTATC AGTTGGACAG TAAGATTTTA TCCCTTTCTC

2281 TTCTGACTGG TATACCTATT TCATTAGTAG CTAGGTGCAC ATATACATCT AGCACAGCTG

2341 TGAGGACAGA CAGAAGGCAA AGTTTCCATG TGGCCTTGAG CAAGTCCCAT CTCACCTCTA

2401 GGCCTCAGTG TCCTCATCTA TAAAATGAGG GACTTCCCTA GAAGTCTTCA TGGTCTCTTC

2461 CAGCCCAGAC ATCCTGTGAT GTCATGAAAG CACCTGCCCT CTGTTTCCCC TCAGAACACC

2521 CTGTACCATC CATGGAGCAC GAGGCCTTCA GAAAAGACAC TTCAATGGGA GTGAACATTT

2581 CTAACTAAGG ACAGGATGGC TGTGTGTGGT GGTCACCAGG TCCTGTGAGC AAAGTGCAGG

2641 TTATGCAAGT CGCCAGGCAG GAGGCCATTC CAGGAGTGGG ATTATTCATC AAACTCTTTG

2701 CCCAGTTCAT CCCAATGGGG GAAGTATTCC CTTCTTTCCT ACTCTGGGAA GAATGTCTCC

2761 TGCCACTCCT CAACTGATGA TAGACTTCGA AAACAGATGA GAAGACTAGC AGCTAGCAAG

2821 GGTGCTTGTA GTCACACTGT GGAACACTAA AGAGCTAGGA AAGAGTTGAG CACAGGCAAC

Figure 5

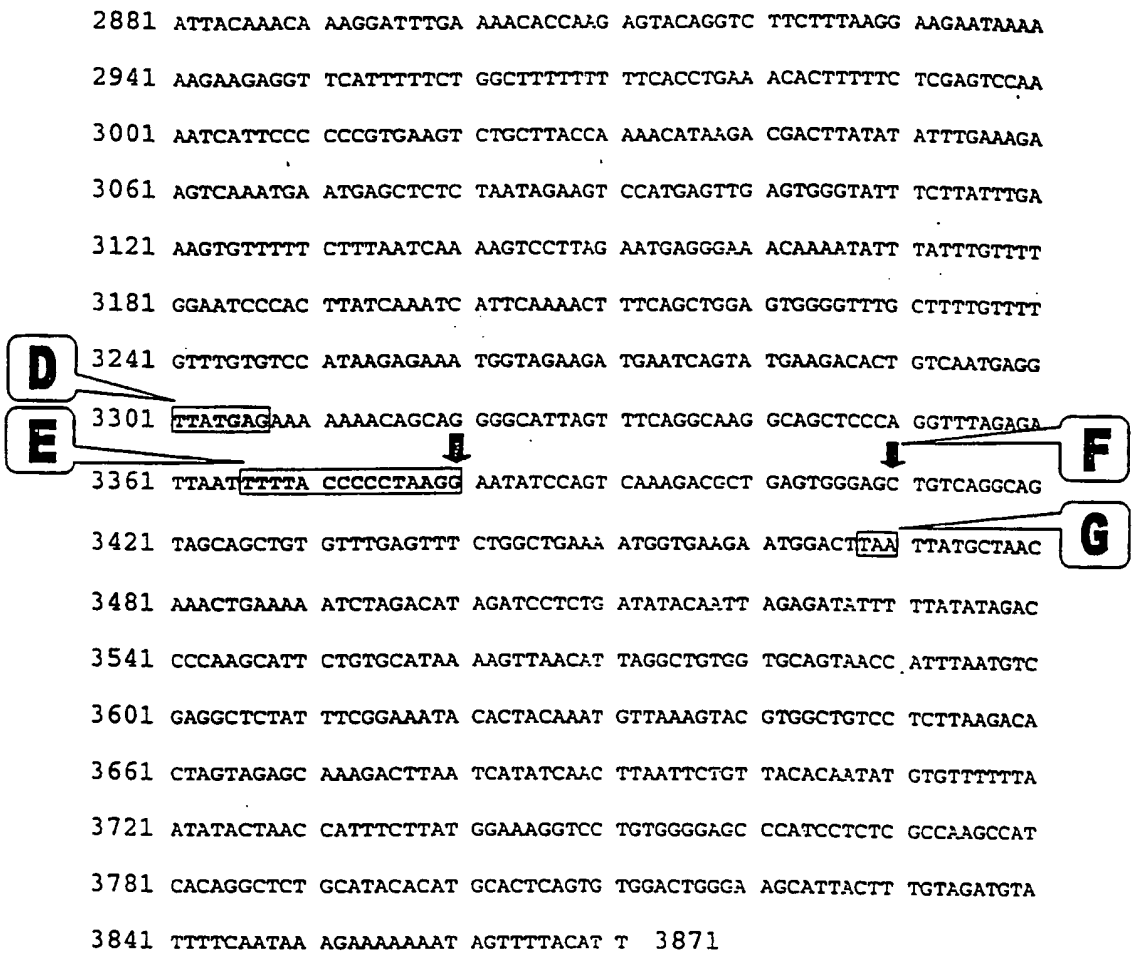

2881 ATTACAAACA AAGGATTTGA AAACACCAAG AGTACAGGTC TTCTTTAAGG AAGAATAAAA
2941 AAGAAGAGGT TCATTTTTCT GGCTTTTTTT TTCACCTGAA ACACTTTTTC TCGAGTCCAA
3001 AATCATTCCC CCCGTGAAGT CTGCTTACCA AAACATAAGA CGACTTATAT ATTTGAAAGA
3061 AGTCAAATGA ATGAGCTCTC TAATAGAAGT CCATGAGTTG AGTGGGTATT TCTTATTTGA
3121 AAGTGTTTTT CTTTAATCAA AAGTCCTTAG AATGAGGGAA ACAAAATATT TATTTGTTTT
3181 GGAATCCCAC TTATCAAATC ATTCAAAACT TTCAGCTGGA GTGGGGTTTG CTTTTGTTTT
3241 GTTTGTGTCC ATAAGAGAAA TGGTAGAAGA TGAATCAGTA TGAAGACACT GTCAATGAGG
3301 TTATGAGAAA AAAACAGCAG GGGCATTAGT TTCAGGCAAG GCAGCTCCCA GGTTTAGAGA
3361 TTAATTTTTA CCCCCTAAGG AATATCCAGT CAAAGACGCT GAGTGGGAGC TGTCAGGCAG
3421 TAGCAGCTGT GTTTGAGTTT CTGGCTGAAA ATGGTGAAGA ATGGACTTAA TTATGCTAAC
3481 AAACTGAAAA ATCTAGACAT AGATCCTCTG ATATACAATT AGAGATATTT TTATATAGAC
3541 CCCAAGCATT CTGTGCATAA AAGTTAACAT TAGGCTGTGG TGCAGTAACC ATTTAATGTC
3601 GAGGCTCTAT TTCGGAAATA CACTACAAAT GTTAAAGTAC GTGGCTGTCC TCTTAAGACA
3661 CTAGTAGAGC AAAGACTTAA TCATATCAAC TTAATTCTGT TACACAATAT GTGTTTTTTA
3721 ATATACTAAC CATTTCTTAT GGAAAGGTCC TGTGGGGAGC CCATCCTCTC GCCAAGCCAT
3781 CACAGGCTCT GCATACACAT GCACTCAGTG TGGACTGGGA AGCATTACTT TGTAGATGTA
3841 TTTTCAATAA AGAAAAAAAT AGTTTTACAT T 3871
D
E
F
G

Figure 6

```
1 ATG ACT AAA AGC AAT GGA GAA GAG CCC AAG ATG GGG GGC AGG ATG
  Met Thr Lys Ser Asn Gly Glu Glu Pro Lys Met Gly Gly Arg Met
   1               5                  10                  15

46 GAG AGA TTC CAG CAG GGA GTC CGT AAA CGC ACA CTT TTG GCC AAG
   Glu Arg Phe Gln Gln Gly Val Arg Lys Arg Thr Leu Leu Ala Lys
                  20                  25                  30

91 AAG AAA GTG CAG AAC ATT ACA AAG GAG GAT GTT AAA AGT TAC CTG
   Lys Lys Val Gln Asn Ile Thr Lys Glu Asp Val Lys Ser Tyr Leu
                  35                  40                  45

136 TTT CGG AAT GCT TTT GTG CTG CTC ACA GTC ACC GCT GTC ATT GTG
    Phe Arg Asn Ala Phe Val Leu Leu Thr Val Thr Ala Val Ile Val
                  50                  55                  60

181 GGT ACA ATC CTT GGA TTT ACC CTC CGA CCA TAC AGA ATG AGC TAC
    Gly Thr Ile Leu Gly Phe Thr Leu Arg Pro Tyr Arg Met Ser Tyr
                  65                  70                  75

226 CGG GAA GTC AAG TAC TTC TCC TTT CCT GGG GAA CTT CTG ATG AGG
    Arg Glu Val Lys Tyr Phe Ser Phe Pro Gly Glu Leu Leu Met Arg
                  80                  85                  90

271 ATG TTA CAG ATG CTG GTC TTA CCA CTT ATC ATC TCC AGT CTT GTC
    Met Leu Gln Met Leu Val Leu Pro Leu Ile Ile Ser Ser Leu Val
                  95                  100                 105

316 ACA GGA ATG GCG GCG CTA GAT AGT AAG GCA TCA GGG AAG ATG GGA
    Thr Gly Met Ala Ala Leu Asp Ser Lys Ala Ser Gly Lys Met Gly
                  110                 115                 120
```

Figure 6

```
361 ATG CGA GCT GTA GTC TAT TAT ATG ACT ACC ACC ATC ATT GCT GTG
    Met Arg Ala Val Val Tyr Tyr Met Thr Thr Thr Ile Ile Ala Val
                    125                 130                 135

406 GTG ATT GGC ATA ATC ATT GTC ATC ATC ATC CAT CCT GGG AAG GGC
    Val Ile Gly Ile Ile Ile Val Ile Ile Ile His Pro Gly Lys Gly
                    140                 145                 150

451 ACA AAG GAA AAC ATG CAC AGA GAA GGC AAA ATT GTA CGA GTG ACA
    Thr Lys Glu Asn Met His Arg Glu Gly Lys Ile Val Arg Val Thr
                    155                 160                 165

496 GCT GCA GAT GCC TTC CTG GAC TTG ATC AGG AAC ATG TTC CCT CCA
    Ala Ala Asp Ala Phe Leu Asp Leu Ile Arg Asn Met Phe Pro Pro
                    170                 175                 180

541 AAT CTG GTA GAA GCC TGC TTT AAA CAG TTT AAA ACC AAC TAT GAG
    Asn Leu Val Glu Ala Cys Phe Lys Gln Phe Lys Thr Asn Tyr Glu
                    185                 190                 195

586 AAG AGA AGC TTT AAA GTG CCC ATC CAG GCC AAC GAA ACG CTT GTG
    Lys Arg Ser Phe Lys Val Pro Ile Gln Ala Asn Glu Thr Leu Val
                    200                 205                 210

631 GGT GCT GTG ATA AAC AAT GTG TCT GAG GCC ATG GAG ACT CTT ACC
    Gly Ala Val Ile Asn Asn Val Ser Glu Ala Met Glu Thr Leu Thr
                    215                 220                 225

676 CGA ATC ACA GAG GAG CTG GTC CCA GTT CCA GGA TCT GTG AAT GGA
    Arg Ile Thr Glu Glu Leu Val Pro Val Pro Gly Ser Val Asn Gly
                    230                 235                 240
```

Figure6

```
721 GTC AAT GCC CTG GGT CTA GTT GTC TTC TCC ATG TGC TTC GGT TTT
    Val Asn Ala Leu Gly Leu Val Val Phe Ser Met Cys Phe Gly Phe
                    245                 250                 255

766 GTG ATT GGA AAC ATG AAG GAA CAG GGG CAG GCC CTG AGA GAG TTC
    Val Ile Gly Asn Met Lys Glu Gln Gly Gln Ala Leu Arg Glu Phe
                    260                 265                 270

811 TTT GAT TCT CTT AAC GAA GCC ATC ATG AGA CTG GTA GCA GTA ATA
    Phe Asp Ser Leu Asn Glu Ala Ile Met Arg Leu Val Ala Val Ile
                    275                 280                 285
                 (A)                              (B)
856 ATG TGG AAT ATC CAG TCA AAG ACG CTG AGT GGG AGC TGT CAG GCA
    Met Trp Asn Ile Gln Ser Lys Thr Leu Ser Gly Ser Cys Gln Ala
                    290                 295                 300

901 GTA GCA GCT GTG TTT GAG TTT CTG GCT GAA AAT GGT GAA GAA TGG
    Val Ala Ala Val Phe Glu Phe Leu Ala Glu Asn Gly Glu Glu Trp
                    305                 310                 315

946 ACT TAA 951
    Thr *
    316
```

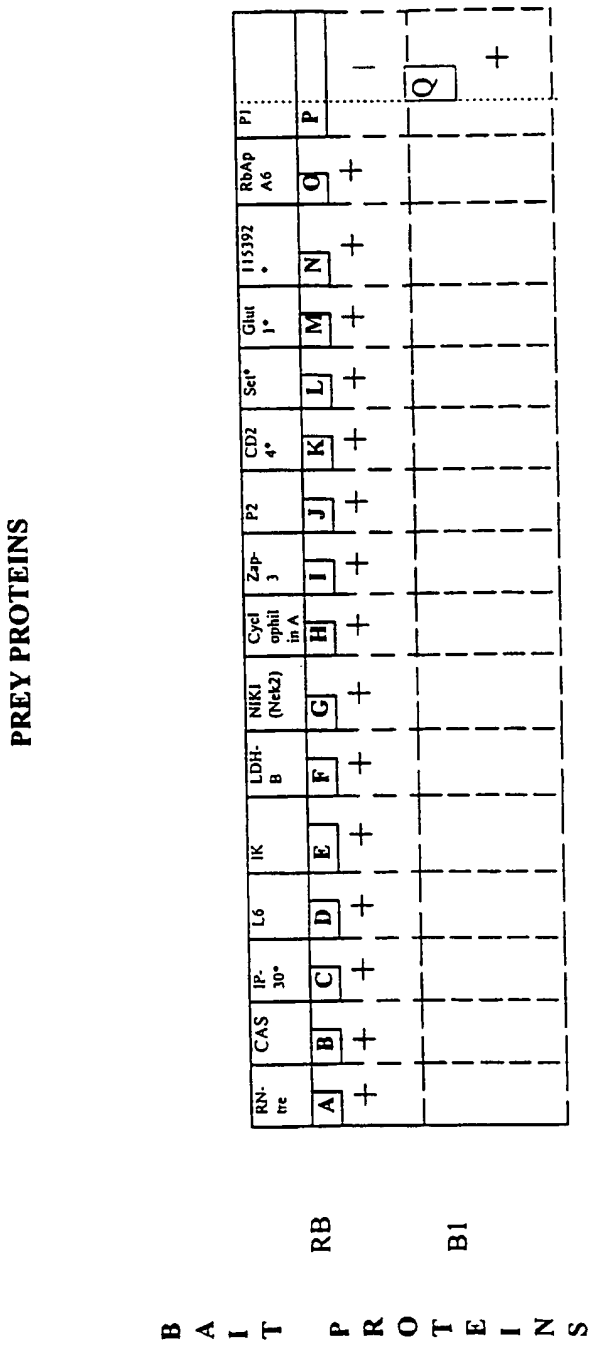
Figure 7

```
ATGACCTTTC CTCTTTATCT TCCTTGTTGT GCAGGTAAAG AAACCAAGTG GAAGAGTGTT     60
TCCTCCTCTG GCCGTAAAGC AGCTGTCCCC GCCCTACTCC GGACCGCCCC AAAGACTCCA    120
TGGGATGGAC CTGAGTCAGC CGAATCCTAG CCCCTTCCCT TGGGCCTGCT GTGGTGCTCG    180
ACATCAGTGA CAGACGGAAG CAGCAGACCA TCAAGGCTAC GGGAGGCCCG GGGCGCTTGC    240
GAAG ATG AAG TTT GGC TGC CTC TCC TTC CGG CAG CCT TAT GCT GGC TTT      289
     Met Lys Phe Gly Cys Leu Ser Phe Arg Gln Pro Tyr Ala Gly Phe
      1               5                  10                  15

GTC TTA AAT GGA ATC AAG ACT GTG GAG ACG CGC TGG CGT CCT CTG CTG      337
Val Leu Asn Gly Ile Lys Thr Val Glu Thr Arg Trp Arg Pro Leu Leu
                 20                  25                  30

                 A
AGC AGC CAG CGG AAC TGT ACC ATC GCC GTC CAC ATT GCT CAC AGG GAC      385
Ser Ser Gln Arg Asn Cys Thr Ile Ala Val His Ile Ala His Arg Asp
            35                  40                  45

TGG GAA GGC GAT GCC TGT CGG GAG CTG CTG GTG GAG AGA CTC GGG ATG      433
Trp Glu Gly Asp Ala Cys Arg Glu Leu Leu Val Glu Arg Leu Gly Met
        50                  55                  60

ACT CCT GCT CAG ATT CAG GCC TTG CTC AGG AAA GGG GAA AAG TTT GGT      481
Thr Pro Ala Gln Ile Gln Ala Leu Leu Arg Lys Gly Glu Lys Phe Gly
    65                  70                  75

CGA GGA GTG ATA GCG GGA CTC GTT GAC ATT GGG GAA ACT TTG CAA TGC      529
Arg Gly Val Ile Ala Gly Leu Val Asp Ile Gly Glu Thr Leu Gln Cys
80                  85                  90                  95

CCC GAA GAC TTA ACT CCC GAT GAG GTT GTG GAA CTA GAA AAT CAA GCT      577
Pro Glu Asp Leu Thr Pro Asp Glu Val Val Glu Leu Glu Asn Gln Ala
                100                 105                 110

GCA CTG ACC AAC CTG AAG CAG AAG TAC CTG ACT GTG ATT TCA AAC CCC      625
Ala Leu Thr Asn Leu Lys Gln Lys Tyr Leu Thr Val Ile Ser Asn Pro
            115                 120                 125

AGG TGG TTA CTG GAG CCC ATA CCT AGG AAA GGA GGC AAG GAT GTA TTC      673
Arg Trp Leu Leu Glu Pro Ile Pro Arg Lys Gly Gly Lys Asp Val Phe
        130                 135                 140
```

Figure 7

```
CAG GTA GAC ATC CCA GAG CAC CTG ATC CCT TTG GGG CAT GAA GTG TGACA   723
Gln Val Asp Ile Pro Glu His Leu Ile Pro Leu Gly His Glu Val
    145                 150                 155         158

AGTGTGGGCT CCTGAAAGGA ATGTTCCAGA GAAACCAGCT AAATCATGGC ACCTTCAATT   783
TGCCATCGTG ACGCAGACCT GTATAAATTA GGTTAAAGAT GAATTTCCAC TGCTTTGGAG   843
AGTCCCACCC ACTAAGCACT GTGCATGTAA ACAGGTTCCT TTGCTCAGAT GAAGGAAGTA   903
GGGGGTGGGG CTTTCCTTGT GTGATGCCTC CTTAGGCACA CAGGCAATGT CTCAAGTACT   963
TTGACCTTAG GGTAGAAGGC AAAGCTGCCA GTAAATGTCT CAGCATTGCT GCTAATTTTG  1023
GTCCTGCTAG TTTCTGGATT GTACAAATAA ATGTGTTGTA GATG                    1067
```

Figure 8


PREY PROTEINS
RN-tre
CAS
IP-30*
L6
IK
LDH-B
NIK1 (Nek2)
Cyclophilin A
Zap-3
P2
CD2 4*
Sel*
Glut 1*
115392 *
RbAp A6
P1
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
+
−
BAIT PROTEINS
RB
B1